Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 39, NUMBER 6 JUNE, 1949

The Production of Diffraction Gratings

I. Development of the Ruling Art*

GEORGE R. HARRISON
Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received February 23, 1949)

In this first paper of several on the design and production of diffraction gratings of high resolving power, an outline is given of the development and present status of mechanically controlled ruling engines. Gratings having notable performance are described, and problems involved in producing on a reflecting surface large numbers of parallel equidistant grooves of constant or controlled form, with position tolerances of one micro-inch or better, are analyzed. Distinctive features of the engines of Rowland, Michelson, Gale, Anderson, Babcock, Siegbahn, and others are discussed. It is pointed out that the low productivity of many ruling engines in the past should not discourage present-day experimenters, as improved surfaces for ruling, superior ruling tools, alloys less susceptible to wear, warpage, creep, and fatigue, better methods of lubrication, and servo-mechanisms capable or making a ruling engine self-correcting, are all now available. Much can be learned by studying the strong points and weaknesses of previous engine designs. A current trend in the use of gratings appears to be away from the concave and toward the plane form. Succeeding papers will deal with a suggested method for obtaining improved gratings having high resolving power, and with the modification and testing of a 14-inch ruling engine, originally designed and partially built at the University of Chicago by Gale and O'Donnell, and recently transferred to M.I.T.

INTRODUCTION

NO single tool has contributed more to the progress of modern physics than the diffraction grating, especially in its reflecting form. The availability of gratings having from 60,000 to 180,000 grooves on an optical reflecting surface of approximately six inches width has made possible the great contributions of line and band spectroscopy to our knowledge of the structures of atoms and molecules. Several hundred gratings having resolving powers ranging from 100,000 to 400,000 have been ruled during the past 65 years, but the demand for gratings still far exceeds the supply. Indeed, active demand is itself apparently discouraged by the limited availability of gratings. Not only do scientists need many more gratings of types already produced, but progress in several fields of optical spectroscopy is being delayed by the lack of gratings of still higher resolving power. For work on the Zeeman effect, on hyperfine and nuclear structures of spectral lines, and on the classification of atomic and molecular spectra, spectrographs capable of covering broad spectral regions in a single exposure with resolution of 400,000 and over are urgently needed.

The resolving power of a grating depends fundamentally on its ruled width, and only secondarily on the number of grooves it contains. Early diffraction gratings as produced by Rittenhouse[1] (1785, 53 slits in ½ inch), Fraunhofer[2] (1823, 4000 grooves in ½ inch), and Nobert[3] (1846, 6000 grooves in 1 inch) were weak in light intensity and low in resolution. The transmission grating has been a teaching rather than a research tool. L. M. Rutherfurd,[4] by ruling as many as 35,000 grooves with a

* Invited paper presented before the New York meeting of the Optical Society of America, March 11, 1949.

[1] T. D. Cope, J. Franklin Inst. **214**, 99 (1932).
[2] J. Fraunhofer, Ann. d. Physik **74**, 337 (1823).
[3] Herr Nobert, whose gratings served for the experiments of Lord Rayleigh and others, kept the details of his ruling process secret. See John Mayall, Jr., J. Soc. Arts, May 8, 1885.
[4] The gratings of L. M. Rutherfurd (*ca.* 1870) are frequently referred to by Rowland, Michelson, and others, but without citation of published works. Rutherfurd was a lawyer interested in astronomy as a hobby; an observatory at Columbia University is named after him. See C. Peirce, Am. J. Math. **2** (1879), and A. M. Meyer, Appleton's *American Cyclopedia*, "Spectrum," **15** (1876).

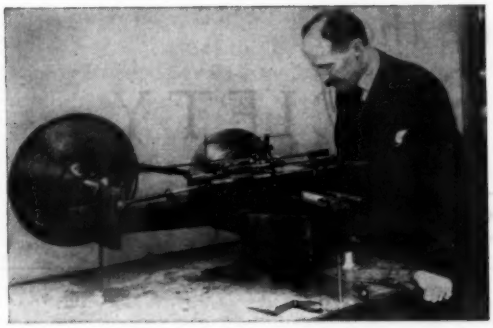

FIG. 1. H. A. Rowland with one of his ruling engines at Johns Hopkins University.

diamond on a speculum mirror 2 in. wide in 1870, was first to produce a grating having resolving power equal to that of available prisms. Many ruling engines of limited success were constructed before 1890, of which that of Mallock[4] is typical.

Father of the modern diffraction grating was H. A. Rowland,[5] who within two years of accepting the challenge of grating ruling had produced gratings of 6-in. ruled width, having more than 100,000 grooves and resolving powers in excess of 150,000. A well-known photograph of Rowland leaning over one of his ruling engines is reproduced in Fig. 1. With his mechanician Schneider, Rowland constructed three engines at Johns Hopkins University which are still in use there; these have produced most of the large gratings of the world. After Rowland's death in 1901 the machines were operated until the Baltimore fire of 1904, in which they were damaged by water, and even more in the repair process. They were finally put in good condition and operated between 1909 and 1916 by J. A. Anderson, who ruled many concave gratings that have never been surpassed, some of which exceed 350,000 in resolving power.

After Anderson's departure the Rowland engines languished until 1923, when their operation was taken over by R. W. Wood. Wood introduced many improvements, and in 1930 began the training of an assistant, Wilbur Perry, who still operates the engines. Wood made the ruling project self-supporting, and supplied scientists throughout the world with many excellent gratings. He continued in charge until several years after his official retirement, when in 1946 direction of ruling operations was taken over by a committee consisting of G. H. Dieke, John Strong, and Wood. The great experimental genius of Rowland and the amazing contributions of his successors have made Johns Hopkins the most dependable source of gratings, and a mecca for those desiring to learn the techniques of grating ruling. Among the outstanding gratings produced at Hopkins are a 6-inch concave ruled by Anderson for the University of Tübingen, with resolving power said to have approached 400,000, and a 7½-inch plane grating recently ruled for the Yerkes Observatory. This last represents the widest ruling that can be produced on the Rowland engines.

A. A. Michelson[6] was the first to accomplish the ruling of a grating considerably wider than 6 in. In 1900 he and S. W. Stratton began construction at the University of Chicago of an engine having capacity for 14 inches of ruled width, shown in Fig. 2. By 1906, when Fred Pearson joined Lange, de Khotinsky, and his brother Julius Pearson in the group assisting Michelson, the machine was producing trial rulings of great promise, and by 1915 it had ruled 8-inch plane gratings with resolving powers approaching 400,000, and one or more 9.3-inch gratings of 110,000 grooves, said to have reached 600,000.[7] Recent inquiry as to the ultimate fate of these outstanding gratings revealed that one had been ruled on a blank too thin to hold its optical figure, and is now a museum piece, and the other was broken soon after completion.

Though a few good gratings were ruled with Michelson's engine, more time was spent in efforts to improve ruling performance than in actual grating production. Late in 1930 H. G. Gale took over direction of the engine and ruled about a dozen 6-in. concave gratings with it. The engine was torn down in 1937, and its central casting was used as the nucleus of an engine of radically new design.

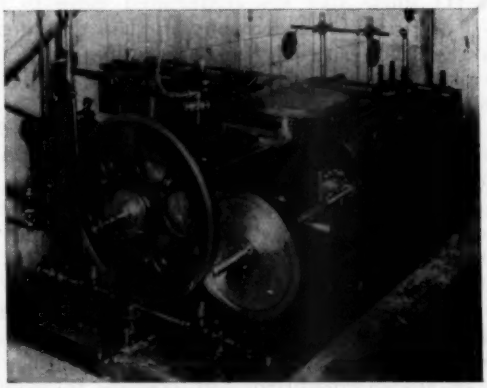

FIG. 2. Michelson's first ruling engine at the University of Chicago.

[5] H A. Rowland, *Physical Papers* (Johns Hopkins Press, Baltimore, 1902); Phil. Mag. **13**, 469 (1882); Nature **26**, 211 (1882); Phil. Mag. **16**, 297 (1883).

[6] A. A. Michelson, Nature, **362** (January 11, 1912).

[7] A. A. Michelson, Univ. Chicago Mag. **7**, 141 (1915); Proc. Am. Phys. Soc. **54**, 137 (1915); J. Franklin Inst. **181**, 785 (1916); *Studies in Optics* (University of Chicago Press, Chicago, 1927), p. 99.

When about two-thirds complete, work on this engine was discontinued in 1942, and at the end of 1947 it was presented to the Massachusetts Institute of Technology, where it is undergoing rebuilding.

Michelson had in 1910 started construction of a second engine, which was completed by 1917 but which still was not functioning satisfactorily when Gale took charge in 1930. This, which will be referred to as the Gale[8] engine, was completely rebuilt, principally with the object of reducing carriage friction, and by 1940 had ruled a dozen or more good 6-inch gratings. Some of these were characterized by high resolving power (300,000 to 400,000) when used to photograph band spectra in high orders, although one that I examined showed undesirable amounts of scattered light around heavy lines when normal exposures were made. Ruling with this engine ceased in 1940, and it was transferred in 1947 to the Bausch & Lomb Optical Company, where it is now being restored to operation.

On leaving Johns Hopkins in 1916, J. A. Anderson went to the Mt. Wilson Observatory and there undertook completion of a very large ruling engine, begun several years before by Jacomini and designed to produce gratings up to 20 inches in width. This engine has thus far produced useful gratings up to only 6 inches in width, 110,000 grooves giving resolving powers up to 160,000. Anderson's contributions to the art and science of grating ruling have been outstanding. His article on ruling in Glazebrook's *Dictionary of Applied Physics* is one of the most useful available.[9]

In 1930 H. D. Babcock,[10] with his instrument maker Prall, began construction of a smaller Mt. Wilson engine, having ruling capacity of 10 in. This engine, shown in Fig. 3, has produced a number of useful gratings, most of which have been plane, as required for astronomical use with telescopes and Schmidt cameras in the Mt. Wilson and Palomar programs. Babcock reports one batch of five plane 7-inch gratings having 10,000 lines per inch, in which the best gratings show resolving powers of about 350,000 in the 7th order from 69,000 grooves.

These three schools of grating ruling, Johns Hopkins, Chicago, and Mt. Wilson, comprise science's only sources of 6-inch and larger gratings up to the present. However, many smaller engines have been constructed. M. Siegbahn[11] built a small

FIG. 3. A view of Babcock's ruling engine at the Mt. Wilson Observatory laboratory in Pasadena.

engine in Uppsala in 1929, capable of ruling 1800 lines in 3 mm of ruled width for x-ray spectroscopy, and later in Stockholm he completed a 4-inch engine that has ruled many excellent gratings, most of which are used in vacuum spectroscopy. This engine is shown in Fig. 4. A larger engine designed by Siegbahn, of ruling capacity 10 in., was started in 1937, is being tested at present, and should soon come into production.

In England Lord Blythswood worked for nearly thirty years on an engine with 12-in. screw, which was transferred in 1908 to the National Physical Laboratory in Teddington. This machine has been described in great detail by Scoble.[12] Few useful gratings were ruled with it before 1920, but between then and 1939 some 250 satisfactory gratings, up to $3\frac{1}{2}$ in. in ruled width, were distributed throughout the world from it by Adam Hilger, Ltd. Only about a dozen gratings have been ruled on the engine since 1939, and these, though sufficiently good for some industrial applications, are not equal to those ruled previously.

Another interesting ruling project was that of H. J. Grayson,[13] a nurseryman of Melbourne, Australia, who in 1893 built a "micro-ruling" engine with which he ruled as many as 120,000 grooves/in., over widths of a fraction of a millimeter, for testing microscope resolution. Later he built an engine for ruling diffraction gratings which is said to have produced usable gratings. Several important suggestions for improving ruling engines were made by Stone,[14] who followed Grayson. Grayson's engine was taken over by Sir Thomas Lyle, who in 1944 presented it to the National

[8] H. G. Gale, J. Sci. Inst. 12 (Zeeman Memorial Volume), 32 (1935).

[9] J. A. Anderson, *Glazebrook's Dictionary of Applied Physics*, "The manufacture and testing of diffraction gratings," p. 30; also J. Opt. Soc. Am. 6, 434 (1922).

[10] Harold D. Babcock, J. Opt. Soc. Am. 34, 1 (1944) and personal communications.

[11] M. Siegbahn, Zeits. f. Physik 62, 435 (1930); Kosmos 8, 186 (1930); M. Siegbahn and T. Magnusson, Zeits. f. Physik 95, 133 (1935); M. Siegbahn, *K. Svenska Vetens Kapsa Kademiens Årbok*, (1940), p. 151.

[12] W. A. Scoble, Collected Researches National Phys. Lab. 8, 213 (1912).

[13] H. J. Grayson, Proc. Roy. Soc. Victoria 30, 44 (1917); W. Stone, J. Sci. Inst. 11, 1 (1934); 14, 8 (1937).

[14] W. Stone, J. Sci. Inst. 11, 241 (1934); *ibid.* 14, 309 (1937).

Standards Laboratory in Sydney, where it is now being put again into operation.

Many hundreds of small gratings of 2½ in. ruled width, distributed in the spectrographs of the Applied Research Laboratories of Pasadena, California, have been produced by Julius Pearson with two engines built and operated by him. These gratings are on speculum metal, and have 24,400 or 36,600 grooves per inch. Though small, they are more effective than prisms of similar aperture. A.R.L. also operate two other engines under the direction of Bruce Bowslaugh, which produce small gratings.

Many other ruling engine projects have been undertaken in past years, and the number of engines that may well operate successfully is increasing. David W. Mann of Lincoln, Massachusetts, is now building two 4-inch ruling engines for the Bausch & Lomb Optical Company. Edward Finn of Springfield, Vermont, has constructed five small engines during the past seventeen years, later models of which show several interesting features, but have not yet produced useful gratings. Edward Leibhardt of Maywood, Illinois, has an engine now well under way. W. W. Hansen of Stanford University is constructing an engine of unique design, the principle of which will be discussed in a later paper of this series dealing with interferometric control of grating ruling.

Optical firms wishing to produce gratings of moderate resolving power show an increasing tendency to copy Rowland's engines with only minor changes of design. Baird Associates of Cambridge, Mass., have recently brought such an engine into successful operation on 4-in. gratings, and are now building a second. The Jarrell-Ash Company of Boston have a somewhat similar engine under construction.

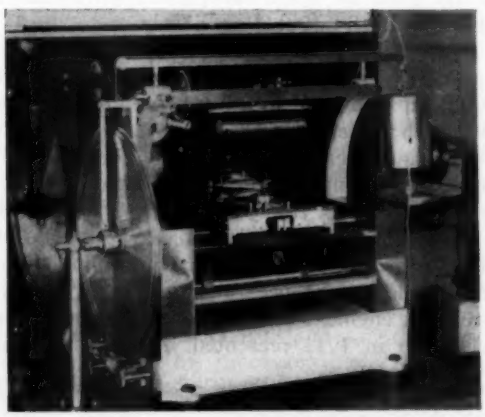

FIG. 4. Siegbahn's ruling engine at the Nobel Institute in Stockholm.

Both Rowland and Michelson emphasized that two-inch gratings are not difficult to rule, but that difficulties accumulate rapidly as ruled width increases. The most successful engines are the small and compact ones patterned on those of Rowland; in the larger engines many new difficulties arise from the rapid increase of weight relative to contact area. In general, success is greater if an engine is kept as small as is consistent with the maximum ruled width desired.

We may sum up the present status of grating ruling as follows: 2-in. gratings of resolving power up to 60,000 can be produced almost on demand on any one of a number of engines, and a suitable new engine of this size could be built and put into operation in less than one year; 4-in. gratings of resolving power up to 100,000 can be produced on at least five engines (Hopkins, Mt. Wilson and Palomar [2], Siegbahn, Baird Associates) with Bausch & Lomb and Jarrell-Ash engines expected to operate within a year; 6-in. gratings of resolving power 150,000 or more can be produced at Hopkins fairly regularly, at Mt. Wilson occasionally, with Siegbahn and Bausch & Lomb probably coming in soon as sources, and perhaps Baird and Jarrell-Ash. Gratings of width 8 in. or more, with resolving powers in excess of 400,000, are not at present available, but there is some hope of production by Siegbahn, and possibly by Bausch & Lomb, by M.I.T., and by Hopkins.

Anderson has pointed out that it is possible to design a ruling engine which theoretically should rule grooves spaced with ten times the precision required for excellent gratings up to 10 in. in width. Unless the elasticity of the materials used to construct the engine is taken into account, however, the attained precision is likely to be less than one-tenth of that required. This hundred-fold range is a "never-never land" of elastic deformation, vibration, creep, warpage, variable lubrication, corrosion, and wear that requires the use of much art as well as science for the successful operation of large ruling engines.[15]

In view of the limited success achieved in the past by outstanding experimental physicists in ruling large gratings, what likelihood does the future hold of better gratings becoming available? We have many advantages unknown to Rowland or Michelson. First is the character of the surface to be ruled; we have soft aluminum[16] deposited on glass instead of the speculum metal they used,

[15] To quote Michelson: "When the accumulation of difficulties seems to be insurmountable, a perfect grating is produced, the problem is considered solved, and the event celebrated with much rejoicing, only to find the next trial a failure···. One comes to regard the machine as having a personality···requiring humoring, coaxing, cajoling, even threatening!" Nature, **362** (January 11, 1912).

[16] John Strong, Phys. Rev. **43**, 498 (1933); Astrophys. J. **83**, 401 (1936).

which is hard and brittle, and hence more suitable for optical polishing than for ruling. A single diamond edge will now often serve to rule several gratings with little variation in groove form, whereas previously a diamond was likely to wear out before half of one large speculum grating was ruled. Second, we have improved methods for the shaping of ruling diamonds. Third, we have improved materials for engine construction, so that castings can be made, normalized, and finished in such a way as to suffer much less from warpage, creep, corrosion, wear, and elastic deformation than formerly. Fourth, we have a better understanding of lubrication and its control, and a wider assortment of suitable dry materials for sliding contacts. Fifth, we have improved electronic devices for the interferometric control of positioning, and servomechanisms that can be used for self-adjustment of a machine. Finally, though early builders of ruling engines were likely to keep their methods secret, we can now profit by the experience obtained in the testing of more than a dozen engines.**

THE DIFFRACTION GRATING

The theoretical resolving power of a grating, R, $=\lambda/d\lambda=Nm$, where $d\lambda$ is the minimum wave-length difference between two lines of wave-length λ that it can just resolve, N is the number of grooves in the grating, and m is the order of interference used. Michelson showed, from the above relation and the well-known grating formula, $m\lambda=(W/N)(\sin\alpha\pm\sin\beta)$ in which W is the ruled width of the grating, that the maximum resolving power attainable with any grating is twice the number of wave-lengths in its ruled width, or $R_{\max}=2W/\lambda$. Thus the maximum theoretical resolving power of a 6-in. grating is 600,000 for light of wave-length 5000A, regardless of how many grooves are ruled on it; the coarser the ruling, the higher the order that is available to achieve the theoretical resolution. Resolving powers of present-day spectrographs could theoretically be almost doubled without requiring use of wider ruled areas, but this result would require more careful control of groove form and better uniformity throughout the ruling process than has yet been attained.

Lord Rayleigh[17] showed that resolving power will not be greatly reduced if rays that are to interfere constructively are not more than $\lambda/4$ out of phase, and from this deduced that grooves in a grating may be shifted from their proper positions by as much as $\frac{1}{4m}$ of the spacing between grooves. In a grating of 15,000 grooves/in., then, occasional grooves could be displaced by as much as 16 micro-inches if only the first order is to be used, or by 4 micro-inches if the fourth order is to be used.***

Unfortunately, reduction of resolution is not the only effect of displaced grooves, for they may cause false lines that are disturbing. If the displacements are irregular, scattered light and a widening of each spectrum line will result; if periodic, ghosts are produced, whose intensities increase rapidly with order m. Rowland[18] ghosts arise from periodic errors in the screw or its mounting, and appear as pairs of lines on either side of every strong line, frequently as multiple pairs. These ghosts may be regarded as spectra about each line as a "central image," produced by a coarse grating having the pitch of the screw as spacing between grooves. Lyman[19] ghosts are produced by two regular periodicities; they may arise when a thickened portion of a driving belt passes periodically over a slightly eccentric pulley, or by regular dripping of oil on the carriage ways, or by periodic turning of ball bearings.

The intensity of the Rowland ghosts from a particular grating usually increases as the square of the order m, and a poor engine may produce ghosts that in high orders are much stronger than their parent lines. The ghost intensity varies as the square of the amplitude of the periodic error, and much time is usually spent in reducing such errors in a ruling engine to diminish the intensities of the Rowland ghosts in the first order to less than 0.1 percent of their parent lines. Lyman ghosts must be reduced even further, as they are often so far from their parent lines that they may be confused with real lines.

Rowland[18] and Meyer[18] have summarized the relations between intensities of false lines and the errors of positioning that produce them. The formula $G_r=\pi^2m^2(e/a)^2$ gives the ratio of Rowland ghosts of the first order**** to their parent line, where m is the order as before, and e/a is the ratio of the maximum amplitude of periodic groove displacement to normal groove separation. From this it will be seen that in a 20,000 grooves/in. grating classed as "excellent," with ghost intensity-ratios less than 1/1000 in the first order, periodic errors must be held to less than 0.5 micro-inch. For a

** Notable improvements in ruling techniques are now being made by John Strong at Johns Hopkins under a grant from the ONR. Preliminary reports show that the problem is being attacked on a very fundamental level.

[17] Lord Rayleigh, *Scientific Papers*, Vol. 1, 1869 to 1881 (Cambridge University Press, London, 1899), p. 458.

*** The effect of order comes from the angle the reflected beam makes with the zero order.

[18] H. A. Rowland, Phil. Mag. **35**, 397 (1893); Astronomy and Astro-Physics **12**, 129 (1893); See reference 5, p. 525; C. F. Meyer, *The Diffraction of Light, X-rays, and Material Particles* (University of Chicago Press, Chicago, 1934).

[19] T. Lyman, Phys. Rev. **12**, 1 (1901); *ibid.* **16**, 257 (1903); W. F. Meggers and C. C. Kiess, J. Opt. Soc. Am. **6**, 417 (1922); Carl Runge, *ibid.* p. 429; J. A. Anderson, *ibid.* p. 434. Rowland and Lyman ghosts, and design of ruling engines.

**** Distinction must be made between Rowland ghosts *of* various orders, and ghosts *in* various orders.

10,000 grooves/in. grating double the amplitude is allowed. In round numbers the criterion has been established that engines should rule precisely to 1 micro-inch, or 0.1 fringe of green light, or $\lambda/20$.

Though gratings with more than 80,000 grooves/in. have been ruled, ghost intensities are greater as the spacing is finer, and the maximum attainable wave-length is twice the groove separation, so 30,000 grooves/in. is usually considered the maximum practical number. The finer the groove spacing the smaller the number of orders between the zero order and the maximum available order, and hence the greater the uncontrolled intensity in any one order; thus no visible light will be thrown into any order higher than the second in a 30,000 grooves/in. grating. Hansen† has suggested forcing all the light into one order by ruling a grating of spacing sufficiently fine that no light will appear in any order above the first; the zero order could then be reduced in intensity by ruling grooves sufficiently deep that no direct reflection would occur, and one of the first orders could be removed by illuminating the grating from such an angle that this order would pass beyond the limiting angle.

Wood,[20] Anderson and Sparrow,[21] Babcock[10] and others have shown that the distribution of intensity among the various grating orders is more readily controlled as groove spacing is made coarser. Thus with 30,000 grooves/in. the groove dimensions exceed the wave-length of the light only slightly, and the intensity distribution in angle around the grating is determined mainly by diffraction. As the groove size increases the intensity distribution is determined more and more by direct reflection; one can then rule V-shaped grooves with plane sides, one shallow and one steep, to direct light by specular reflection so that the intensity in all orders falling within a fairly limited angle will be increased. Wood[20] carried this to the extreme when he devised the coarse echelette grating, with from 1000 to 2000 lines per inch, in which almost all the light is reflected specularly from one flat side of the groove. The resulting "blaze" falls within a very narrow angle, so that only two or three close orders receive all the light. The echelette has been found especially useful in the infra-red region. Randall[22] has described the ruling of echelettes on a large engine at the University of Michigan. Spacings ranging from 25 to 5000 grooves/in. are used. Since the waves to be dispersed are many times as long as those of visible light, precision requirements are greatly relaxed in this engine.

Wood,[23] Babcock,[10] and others have extended the earlier work of Rowland and Anderson in intensifying one or more orders of gratings having 15,000 grooves/in. by so shaping and orienting the ruling diamond that more light is thrown to one side than to the other. They have thus succeeded in throwing as much as 60 percent, including reflection losses, of the incoming green light into one order. Even though these gratings are designed for the visible and ultraviolet regions they are sometimes referred to as "echelettes." This usage leads to ambiguities, and it may prove desirable to reserve the term echelette for coarse gratings of controlled groove form designed primarily for infra-red use. Gratings of narrow spacing in which orders on one side have been brightened by controlling groove form might well be called "blazed gratings."

Groove shape can be held more constant during the ruling of a coarse grating than a fine one, which probably explains why higher resolving powers have been reached in the 6th orders of 10,000 grooves/in. gratings than in the 2nd orders of 30,000 grooves/in. gratings. Indeed, many of the latter show little more resolving power in the second order than in the first. Wood,[24] after a suggestion of Shane, has used echelette gratings to disperse visible light in orders from the 15th to the 30th but has not thus attained higher resolving powers than with ordinary gratings. In the second paper of this series some possibilities will be discussed for producing compact grating spectrographs of high resolving power by using very coarse plane gratings of a special type.

Though the concave diffraction grating has come into much wider use than the plane grating, the highest resolving powers have been attained with the latter. The concave grating requires no expensive auxiliary lens, gives broad spectral coverage in a single exposure, and has wide flexibility; it suffers from astigmatism, however, gives less intensity and resolution than the plane grating, and is more difficult to rule.

The trials of ruling a grating blank increase with its concavity for several reasons. The diamond must be pivoted with more free vertical motion when a deep concave surface is to be ruled, and hence is more difficult to stabilize laterally than when a plane grating is ruled. "Chatter," a vertical oscillation of the ruling tool that produces periodic variations in groove depth, is more troublesome with concave surfaces than with plane. Again, side thrust on the ruling tool varies with the slope of the surface that is being ruled. Most diamonds rule by pressing metal aside rather than by actually

† Private communication.

[20] R. W. Wood, Phil. Mag. **4**, 396 (1902); *ibid.* **12**, 585 (1906); *ibid.* **14**, 477 (1907); *ibid.* **20**, 770 (1910); *ibid.* **23**, 310 (1912): A. Trowbridge and R. W. Wood, Phil. Mag. **20**, 886 and 898 (1910).

[21] J. A. Anderson and C. M. Sparrow, Astrophys. J. **33**, 338 (1911).

[22] H. M. Randall, J. App. Phys. **10**, 768 (1939).

[23] R. W. Wood, J. Opt. Soc. Am. **34**, 509 (1944).

[24] R. W. Wood, J. Opt. Soc. Am. **36**, 715 (1946).

cutting it, and are likely to rule with different parts of the edge on the uphill cut from those used when going downhill; a change in groove shape then results which produces "target pattern"[25] and lowers resolving power, especially in the higher orders.

Plane gratings are ordinarily ruled with grooves ½ to ¾ as long as the grating width, to give a square aspect to the ruled surface at the angle of incidence most frequently used in spectrographs. Concave gratings of even 10 meters radius of curvature are seldom ruled with grooves longer than 2 in. because astigmatism increases as the groove is lengthened. A plane grating used with a suitable lens is therefore likely to be from two to ten times as fast as a concave grating, because of its greater ruled area and the absence of astigmatism. Only a short spectral range can be photographed in one exposure with a plane grating, however, because of the limited field of the lens used to collimate light on it.

While the concave grating still remains unsurpassed for many purposes, the increasing use of the Schmidt camera and the rising importance of infrared spectroscopy, in which concave mirrors can effectively be used in spectrometers, have combined to increase the importance of the plane reflection grating.

DESIGN OF THE RULING ENGINE

The success of a ruling engine depends on how well compensation has been made for the inevitable changes in shape and size that result from variations in friction, temperature, elasticity, and internal strain. The designer may well imagine his engine as constructed of rubber, thus visualizing the stresses that must be controlled when strains of the order of a micro-inch become of great importance. Present carefully designed engines must be set into operation from 3 to 24 hours before the diamond is lowered into the ruling position, so that thermal and dynamic equilibrium may be attained. Wood[26] found that stopping one of the Rowland engines for only a few minutes produced marked displacement of several grooves when ruling was resumed; the effect presumably arose from changing thicknesses of oil films.

The most detailed descriptions of ruling engines and discussions of their design are those given on behalf of Rowland,[27] by Anderson,[9] and by Scoble.[12] All successful designs have been based on that used by Rowland, in which the grating blank is moved forward the distance between grooves by a screw driven by a ratchet or worm wheel, between strokes of the ruling diamond. To reduce effects arising from intermittent motion of the blank the Baird Associates' engine keeps the blank moving continuously during ruling. The resulting inclination of the grooves to the grating width requires only that the grating be rotated slightly when mounted; however, it is important that the grooves be kept fairly straight and exactly parallel, and this is accomplished by using elliptical gears to connect the motions of the diamond and blank.

A popular misconception exists that the critical portion of a ruling engine is its screw. A screw that will produce translations precise to one micro-inch without need of an auxiliary correcting mechanism can be made readily by the lapping method of Rowland.[28] Michelson[29] states that the time required to lap a screw varies as the amount of material to be removed, and that the lapping of the 20-in. screw for his engine required 15 years. However, I have been informed by T. J. O'Donnell, who supervised the production of the 20-in. screw now in the M.I.T. engine, that this required for its complete production only about 3 months of one man's time, spread over a period of one year. This screw was lapped to within 3 micro-inches periodic error in any turn, and 7 micro-inches "error of run" (departure from uniformity of advance after periodic error has been corrected for) over the entire screw.

Substitutes for the screw suggested by persons who have been over-impressed with the difficulty of making one have thus far demonstrated no superiority. Devices for "inching" the carriage ahead by uniform steps are of value only if they include a means of preserving the location of the first groove rather than that most recently ruled, and their designers usually overlook the requirement that all grooves be located within one micro-inch of their correct positions relative to the first groove, *no matter how many inches away that may be.*

It is important to provide some mechanism such as a cylindrical cam for correcting periodic errors and errors of run, for even when these corrections are not required initially they are likely to be needed later. Even some small engines have been found to require re-correction for periodic error twice a year, and for error of run every few years. Even carefully made screws are likely to warp. A heavy plate carriage was found to have warped in six years so that one end was 0.013 in. away from parallelism with the ways on which it was to move. While modern methods of heat treatment should

[25] R. W. Wood, Phil. Mag. **48**, 497 (1924); George R. Harrison, *Proceedings Seventh Summer Conference on Spectroscopy and its Applications* (John Wiley & Sons, Inc., New York, 1940).

[26] R. W. Wood, Phil. Mag. **48**, 497 (1924).

[27] *Physical Papers of Henry A. Rowland* (Johns Hopkins Press, Baltimore, 1902), p. 691.

[28] Article "Screw," *Encyclopedia Britannica*, 9th Ed., Vol. 21; See reference 5, p. 506.

[29] A. A. Michelson, Nature **88**, 362 (1912); J. Franklin Inst. **181**, 785 (1916).

greatly reduce such warpage, this experience makes it easier to understand why many large engines have stood for years without useful production, other parts warping while each was being corrected.

The body of the ruling engine should be as sturdy and compact as possible. It should be supported on three feet, preferably resting, respectively, in a conical hole, in a V-shaped groove, and on a plane. Rowland obtained equivalent freedom from strains by mounting his engine on three tall spindly legs (Fig. 1), an excellent device for small engines.

Rowland warned against using a screw of pitch so fine as to be difficult to lap, and most engines have screws of 1, 1.5, or 2 mm pitch, or 20 or 30 threads per inch. The screw is made sufficiently longer than the maximum width of ruling desired, to allow for a nut covering from 4 to 6 in. of thread. Thus an averaging effect is obtained by simultaneously contacting from 50 to 150 threads. The diameter of the screw is made sufficiently great that it will sag only a few micro-inches in the center. The screws on Rowland's engines that rule $7\frac{1}{2}$ in. are $1\frac{1}{2}$ in. in diameter, while that on the M.I.T. engine, which is designed to rule 14 in. of width and hence has 20 in. of thread, is ground on a 3-in. tube.

A standard V-thread of 52° angle is ordinarily used for ease in lapping, but T. J. O'Donnell introduced the buttress thread in the Gale and M.I.T. engines. In this, one side of the thread is perpendicular to the screw axis and the other inclined at 45°, the nut making contact on the vertical side. Curvature of the screw or of the carriage ways, and lack of parallelism between these, then produces less undesired carriage translation than when a V-thread is used. The buttress thread has been criticized as difficult to lap because, not being symmetrical, the halves of the nut cannot be reversed during grinding. O'Donnell succeeded in making excellent screws of this type, however, by interconnecting two screws with a carriage so that one guided the lapping of the other.

The screw is mounted in sustaining bearings near each end, and an end-thrust bearing is provided to prevent longitudinal motion of more than a fraction of a micro-inch. The sustaining bearings should be self-aligning, and should not be strained when the length of the screw changes. Journal bearings are likely to introduce periodic errors, and to introduce irregular frictional forces when the screw is large. In the Siegbahn engines the journal bearings are relieved so as to touch the screw in four small areas only. Gale and O'Donnell used highly selected balls rolling in carefully lapped races to mount the screws in the Gale and M.I.T. engines. Great care must be exercised to avoid the introduction of dust particles when bearings are only partially enclosed.

The end-thrust bearing is especially important, since it must hold the screw in position against the frictional forces required to move the carriage. Irregularities in the bearing plate will introduce end motion of the screw having periods commensurate with one turn, producing Rowland ghosts. Wear of the thrust bearing during ruling will introduce other errors, and very hard thrust materials should be used. In some large older engines the end-thrust was as great as 40 lbs., and much attention has been given to reducing this. Gale found that the motion of a heavy carriage under sliding friction was so erratic as to vary from 3 to 10 fringes when the screw was turned to produce an advance of 6 fringes each time. Michelson reduced sliding friction of the carriage by relieving 90 percent of its weight by means of floats operating in mercury, and this method was also used at Mt. Wilson. Gale[8] made use of two carriages, a method described independently by Stone.[14] In this the main load of a slave carriage is driven by a slave screw of low precision, while the precision screw is required to move only the master carriage by the small amount needed to bring this into its proper position. The master carriage is supported by balls rolling on the slave carriage or by steel ribbons hanging from it; even though the balls are required to roll only a few micro-inches they must be highly precise. Much of the weight of the two carriages is relieved by mercury floats.

Siegbahn[30] relieves the ways of much of the carriage weight by mounting an auxiliary screw above the engine. This screw drives a carriage on which are mounted pulleys over which weights are suspended on cords. These pull upward on the carriage. Gale and O'Donnell reduced friction in the M.I.T. engine by carrying the full weight of the carriage on specially selected balls, so that end thrust is reduced to one pound.

For a thrust bearing Rowland[27] used a flat end-plate on the screw, of diamond or other hard material, adjusted perpendicular to the screw axis and bearing against a rounded steel pin, as shown in Fig. 5a. Anderson[9] describes a similar arrangement in which a reconstructed ruby polished flat to $\frac{1}{10}$ fringe is used. He could discern no signs of wear in this after several years of operation against a thrust of 40 lbs. Siegbahn† uses a flat steel end plate that presses against a sapphire hemisphere, adjusted centrally by means of Newton's rings, as in Fig. 5b. In the Gale and M.I.T. engines the end of the screw was lapped flat normal to the axis, and pressed against a thrust-block of babbitt metal under oil. Stone[14] suggested the "screw of zero pitch" shown in Fig. 5c. This is well-thought of by many, and is being tested by Strong. Richardson†

[30] M. Siegbahn, Kosmos 8, 186 (1930).

at Bausch & Lomb has replaced the babbitt metal with Graphitar, a very hard substance of graphite and carbon produced by the United States Graphite Company of Saginaw, Michigan, mounted on a self-aligning plate in accordance with a design by H. Kurtz shown in Fig. 5d. This thrust bearing was adopted for the rebuilt M.I.T. engine, as it eliminates high concentrations of stress.

In the Rowland and Blythswood engines the screw is turned by means of a ratchet wheel having up to 1000 teeth; in most other engines a worm wheel is used. In the Gale and M.I.T. engines, for example, a wheel of 12 in. diameter is mounted on the end of the screw without auxiliary bearings. This worm wheel has 540 teeth, and is driven by a worm which in turn is operated by a ratchet wheel having 40 teeth. It is important that the pawl operating this ratchet attack the tooth without producing appreciable vibration, which introduces Lyman ghosts.

The nut used by Rowland was made by cutting threads in the end grain of *lignum vitae* plugs inserted in a bronze or steel cylinder.[28] Later babbitt metal was substituted for the hard wood, though the latter has much to recommend it from a lubrication standpoint if properly selected. In general the threads of the nut should be of softer material than the screw, and must be protected from dust.

Since it is impossible to mount the axis of the screw exactly parallel to the ways on which the grating carriage travels, the nut should not be rigidly connected with the carriage. Rowland pointed out the impracticability of resting one end of the carriage on the nut, and designed a yoke that permits the nut to move laterally while pushing the carriage ahead. O'Donnell designed a multiple nut which is used in the Gale and M.I.T. engines; the threads of this cover only about one-tenth of the circumference of the screw. Seven separated segments of 8 threads each are mounted on the bottom of the carriage, vertical motion introducing little error because buttress threads are used. Siegbahn reduced friction and improved lubrication by cutting away most of the threads in a solid split nut, so that only four small areas in each half contact the screw.

A counterweight is used to keep the carriage and its nut pressing against the screw thread. In modern engines this weight may be as little as 3 lbs., but if made to drive the carriage it should be sufficient to preload the screw and the thrust bearing and thus reduce elastic deformation, in addition to overcoming carriage friction.

One of the ways on which the grating carriage rides is usually V-shaped and supports the carriage at two points, while the second way is flat, with a single slider. Siegbahn has adopted round ways,

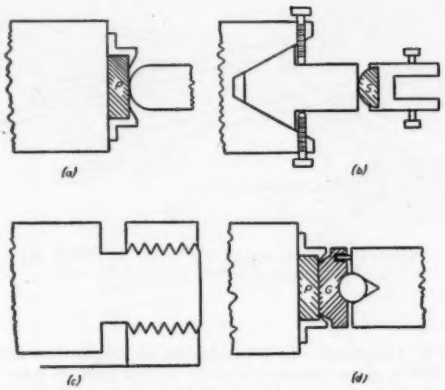

FIG. 5. Typical thrust-bearings for ruling engine screws: (a) A diamond or ruby plate *P* on the end of the screw bears against a spherical steel surface, as used by Rowland and by Anderson; (b) An adjustable steel end-plate bears against a sapphire ball *S*, as used by Siegbahn; (c) Screw of zero pitch, as suggested by W. Stone; (d) Self-aligning thrust bearing with low stress concentration, designed by Henry Kurtz and used under oil in the Bausch & Lomb and M.I.T. engines. *P*, steel or ruby plate; *G*, Graphitar plate.

with two V-shaped sliders and one flat one. O'Donnell used for each way two hard steel rods held closely side by side with steel balls rolling in the channel between them. The ways must be straight and parallel to the axis of the screw. Michelson tested ways by mounting an 8-inch plane mirror on the grating carriage, with a large lens of 30-ft. focal length held stationary in front of this. He then used this auto-collimator to observe with a telescope the position of a pair of cross-hairs as the carriage traversed the ways. He stated that it is necessary to make the ways straight to within 1 second of arc, which is about 50 micro-inches in 10 in. of motion. An alternative method of testing involves laying an optical flat on the carriage and setting up fringes between this and an interferometer plate placed above it; when the two plates have been made parallel to the ways in one position, vertical displacements of the carriage can be measured to 1 micro-inch (0.1 fringe) with ease.

The ways of the ruling carriage, at right angles to the main ways and hence sometimes called the crossways, may also be V-shaped or round. Round ways have the advantage that they can be removed from the engine and made straight and smooth by grinding and lapping in a lathe; flat ways must be scraped in position whenever they warp or wear. Rowland made his V-ways with a 90° included angle, one side being made horizontal to bear the weight of the carriage, and the other vertical to eliminate side motion. Springs pressed the carriage sideways against the vertical faces of the ways. Rowland used ways on both sides of the engine, with lateral constraint in one direction on the front

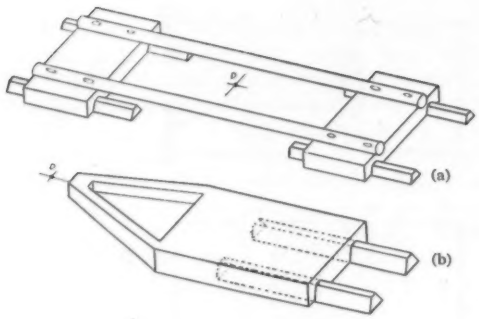

FIG. 6. Diagrammatic presentation of single and double ways. For a given amount of wear in one pair of ways the lateral motion of point *D*, above the ruling tool, is cut in half in (a) but nearly doubled in (b).

pair, and in the other on the rear. Lateral motion of the diamond arising from wear or irregular lubrication is lessened by using two pairs of ways, as shown diagrammatically in Fig. 6a. In the Chicago engines the ruling carriage runs on a single pair of ways set at the rear, as in Fig. 6b. Lateral motion of the diamond is here increased by a factor of almost four over the other arrangement.

A monorail suspension designed by the writer and used for the ruling carriage in the M.I.T. engine is shown in Fig. 7. Two Graphitar sliders are placed 18 in. apart on a 40 in. Nitralloy rod of 1.5 in. diameter, which has been carefully lapped round and straight, and is supported at its midpoint and ends. The frame connecting the sliders carries a ball bearing at each end, from which depend the two ends of the ruling carriage. A boxwood shoe *S* on one side of the carriage bears against a glass optical flat *F* with a very small component of the carriage weight. The optical flat, free from load and wear, can thus exercise almost complete lateral control over the position of the diamond.

Because the ruling carriage ways are likely to wear rapidly and unevenly owing to the high sliding speed and rapid reversals of the carriage, Siegbahn eliminated them by holding the ruling tool in a double-jointed lever. The wide base of this light element, made of a magnesium alloy, rests in sapphire V's.

Strong[31] is investigating the desirability of changing from the "shaper" type of engine hitherto used, to a "planer" type, in which the blank is moved against the tool to form the groove. This has the advantage of reducing the force on the thrust bearing to a few grams. A hydraulic drive is used to operate the reciprocating motion of the blank, and it is hoped that ruling speeds can be greatly increased by this means.

[31] John Strong, ONR Report, Nov. 11, 1948.

THE RULING PROCESS

The introduction by Strong[16] of the aluminum reflecting surface deposited on glass by evaporation, as utilized by Julius Pearson, by Babcock, and by Wood,[32] is one of the most important modern improvements in ruling technique. Not only is the reflecting power of aluminum greater than that of speculum, especially in the ultraviolet, but aluminum gratings show much less scattered light. Speculum gratings can be coated with aluminum or other metals[33] to improve brightness and decrease exposure time, but this increases the intensity of the scattered light proportionally, thus offering no improvement in "signal-to-noise" ratio. Gratings ruled directly on aluminum scatter less light than prism optical systems in many instances.

Of even more importance is the softness of the aluminum coating and the resulting diminution of diamond wear. Speculum metal is so hard that in one instance a diamond that had been used to rule a 15,000 grooves/in. grating, when later used to rule a 5000 grooves/in. grating, nevertheless produced 15,000 grooves/in., having developed a serrated edge in the first ruling. Speculum metal contains large irregular grains that are separated by deposits of softer material, and these sometimes deflect the diamond into a previously ruled groove. Speculum varies considerably from specimen to

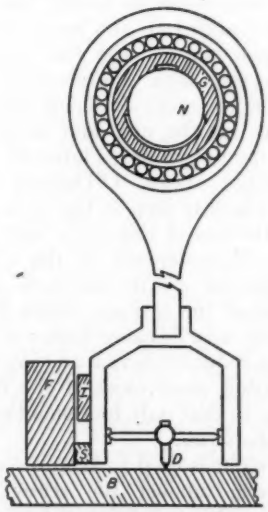

FIG. 7. Monorail construction for ruling carriage as used at M.I.T. Transverse motion of the sliders is reduced 100-fold at the ruling tool *D*, whose lateral position is controlled by the optical flat *F* through the boxwood shoe *S*. *I* and *F* form an interferometer for detecting lateral motion.

[32] The first Hopkins grating on Al went to the Dominion Observatory in 1935; see Reference 63.

[33] E. Gehrke and G. Leithaüser, Ber. d. D. Phys. Ges. **11**, 310 (1909); T. Lyman, Astrophys. J. **60**, 11 (1924); J. Strong, Astronom. Soc. Pac. **46**, 10 (1934).

specimen, which probably accounts for variations in ease of ruling that have been reported. Small pits in a speculum surface are particularly objectionable, as they tend to break down the diamond edge.

Such statements in the literature as "No more speculum blanks being available, it was necessary to use aluminum," and the long delay in certain laboratories in adopting aluminum-coated blanks, probably arose from attempts to rule unsuitable aluminum surfaces. While deposition by evaporation is not difficult to carry out, thorough cleaning of the glass surface, and in some cases use of a chromium undercoat, is needed to insure a film that will adhere permanently after ruling. It is of especial importance to avoid depositing small hard specks in the film. Films prepared by John Strong and those of the Evaporated Metal Films Company of Ithaca, New York, have been found especially suitable for ruling.

Strong has suggested ruling gratings in a film of gold deposited on glass, and this has been done by Babcock.[10] A surface film of aluminum can then be deposited to increase reflecting power. Such gratings can be cleaned by removing the aluminum coating and recoating, whereas gratings ruled in aluminum are somewhat fragile, especially in the presence of hydroxide solutions and halogen vapors. Wood[34] produced a number of gratings by ruling directly on glass, and then etching with hydrofluoric acid to deepen the grooves. For the production of replicas certain grades of stainless steel have been ruled by Babcock and others.

The grating blank, which should be not less than $\frac{1}{8}$ its diameter in thickness, must be mounted firmly on its carriage for ruling, and should be as level as possible. Care must be taken to avoid distortion. Ordinary oil paint is an excellent slow-drying adhesive that does not introduce strains arising from heat or rapid shrinkage.

Authorities differ as to the mechanism of ruling, some claiming that the diamond should "sing" while doing its best work, and others that it must remain silent. Rowland used natural diamonds of about $\frac{1}{3}$ carat for ruling. Anderson developed the "boat-shaped" diamond, which was ground and polished with two curved sides that intersected to give a sharp ruling edge like the keel of a canoe. Others have used "chip" diamonds, which have been found by many to crumble away during the ruling process. Much progress has recently been made in the shaping of diamonds for ruling, especially at the National Bureau of Standards, where a sparking procedure has empirically been found to speed up the shaping.[35]

Various estimates are found in the literature as to the force with which the diamond should bear

[34] R. W. Wood, Phil. Mag. **48**, 507 (1924).
[35] R. W. Wood, Phil. Mag. **20**, 770 (1910).

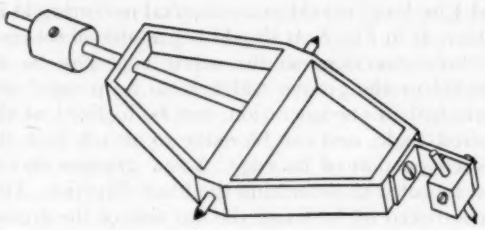

FIG. 8. Mounting of the ruling diamond to give adjustment in three planes.

on the metal. The groove is to be pressed rather than cut, and the pressure needed for this is several tons per square inch of edge in contact. With a chisel-edge diamond a force of from 3 to 5 grams is used at Johns Hopkins for ruling aluminum on glass. For soft materials like Al and Sn small loadings suffice; with speculum or steel a heavier loading is necessary. Babcock recommends coating the plane blank with a thin layer of silicone oil to reduce sticking of material to the diamond, but on most engines no ruling lubricant is used.

The diamond must be lifted from the blank during the return stroke, and then lowered very gently at the beginning of the ruling stroke. Scoble[12] and Siegbahn[30] have described studies showing the importance of bringing the diamond in tangentially to the blank at the beginning of the cut, to avoid bouncing of the tool, pitting, and dulling of the ruling edge. It is usually not sufficient to let the diamond holder ride on a static cam to control this lifting process; a positive motion is needed which can best be introduced by an independently moving cam, operated direct from the main drive shaft.

When ruling a plane grating the diamond is lifted only a few thousandths of an inch on being returned to form the next groove, but as much as 5-mm vertical motion may be needed with a concave grating of short radius. One of the most difficult problems to solve is that of keeping the diamond always in the same plane no matter at what height it is ruling. In the Rowland engines a steel pivot is used which moves in steel bearings. Babcock uses a double spring support which, being very stiff, is most suitable for plane gratings. O'Donnell developed a rolling pivot for use in the Gale engines. Many years ago O'Donnell, and recently David Mann, have used small pre-loaded precision ball-bearings for a very flexible mount.

O'Donnell and others have carried out experiments in which a small interferometer plate about 1 cm square was mounted on the diamond holder, and interference fringes were then produced between this and a large flat plate so that irregular motions of the diamond could be observed.

The diamond, after being set at the proper angle in the end of a short rod about $\frac{1}{8}$ in. in diameter

and 1 in. long, is held in a universal mounting in its holder, as in Fig. 8. It should be capable of rotation in three directions so the ruling edge can be set parallel to the groove which is to be pressed and burnished in one operation, can be inclined at the desired angle, and can be made to attack with the desired portion of its edge. A few grooves should first be ruled to determine the blaze direction. This is controlled by inclining the flat side of the groove so that specular reflection throws the light in the desired direction. If the center of the blaze is such as to make intense the first-order green on one side, the second-order ultraviolet of half wave-length in that same direction will also be intense, unless groove irregularities or other factors intervene. Babcock[10] gives figures which indicate that 60 percent of the light can be thrown into one order with a 15,000 groove/in. grating, 70 percent with 10,000 grooves/in. and 85 percent with 2000, and Wood has reported similar results. Babcock states that several weeks of adjustment and trial ruling are sometimes necessary before a satisfactory grating can be ruled, the final process taking only 70 hours.

ENVIRONMENT AND OPERATION OF THE ENGINE

The Hopkins engines are mounted in rooms measuring about 9×9 ft., while the engine room at M.I.T. is 10×15 ft. The smaller Mt. Wilson engine is in a room of 8×13 ft., but Babcock recommends at least 13×13 ft. The Bausch & Lomb ruling engine room measures 16×30 ft., with facilities for operating three engines simultaneously in the one room. The Hopkins and M.I.T. rooms are lined with aluminum foil to assist in radiation control and to keep down dust.

Most engines are provided with a housing of glass or Lucite immediately surrounding them, which can be kept closed when the engine is in operation, to aid in temperature and dust control. This cover should be easy to open, and readily removable, as it is important to be able to get at all parts of a ruling engine quickly for inspection, cleaning, and overhaul. The inner housing should be maintained at a higher temperature than the outer, and its thermostating should be independent of that of the outer room so that it can be kept constant to within 0.01°C when the outer room is kept constant to within 0.1°. It is especially important that variable heat losses through the base of the machine and the floor be minimized. Where a cork insert cannot be used it is desirable to leave a space for air circulation under the engine. The air in the engine room should be kept well stirred with slow-moving fans.

The engine base is usually mounted on a heavy stone or concrete bed, or on antivibration mounts when this is necessary. Scoble[12] mentions that vibration of trees in a heavy wind ruined the operation of the Blythswood engine; Wood has referred to the effect of passing streetcars on gratings ruled on the Hopkins engines. Bausch & Lomb eliminate vibrations from large ball mills by using antivibration springs under the engine base. Baird Associates, on the other hand, take no special precautions against vibrations, though their Rowland-type engine operates near a car barn. In the absence of quantitative figures, it seems probable that vibrations which produce amplitudes no greater than 1 micro-inch in the plate carriage or diamond tool are not deleterious.

It is desirable that the motor which drives the engine be mounted outside the ruling room, to reduce vibration and thermal effects. The motor is usually operated on d.c., to reduce vibration, so it can be floated on storage batteries to avoid power failure, and to give flexibility of speed. A $\frac{1}{4}$ h.p. motor is suitable for a small engine, but $\frac{3}{4}$ h.p. is useful for a large engine to reduce effects of changes in load. The motor is ordinarily belted to a drive-shaft that extends through the wall into the engine room. This shaft is then coupled, either directly or through a belt drive, to the backshaft of the engine, which carries the cams and eccentrics that advance the grating carriage and give reciprocating motion to the ruling carriage. Most engines operate at from 8 to 20 strokes per minute, so that an average grating can be ruled in 72 hours. A 6-in., 30,000 groove/in. grating usually requires about 12 days to rule.

Lubrication problems of a ruling engine are among the most difficult to solve. Babcock recommends acid-free Vaseline for lubrication of parts requiring grease. Most of the older engines use oil cups on sliders, with felt pad wipers to remove dust. Richardson at Bausch & Lomb has experimented with Graphitar sliders, both dry and with oil lubrication. Dry Graphitar sliding on nitralloy has been used successfully at M.I.T. for the ruling carriage slide. Intermittent oiling during ruling is undesirable, as the oil films involved are much thicker than the 1-micro-inch tolerance allowed, and must be kept as constant as possible.

Gale and O'Donnell designed the engine now at M.I.T. to be immersed in some 350 gals. of oil of medium viscosity, up to within an inch of the top of the grating carriage. Such immersion, besides providing uniform lubrication, gives a thermostatic bath of high heat capacity, aids in distributing heat to give uniform temperatures, and serves as a vibration damper, as well as washing dust from bearing surfaces. It is extremely important to keep this oil well filtered.

The screw, nut, and all moving parts must be kept scrupulously free from dust. Scoble[12] cleaned the screw of the Blythswood engine with benzole and a wire brush after each grating was ruled.

TESTING ENGINE PERFORMANCE

Performance tests that gratings should satisfy have been outlined in the literature.[36] The operator of a ruling engine should be able to estimate its performance quantitatively in terms of certain of these tests. In particular the Rowland ghosts should in the first order be reduced to less than 1 part in a thousand of the intensity of their parent line. This requires elimination of periodic error greater than 1 percent of the grating space.

Many methods of estimating periodic error have been used. Rayleigh[37] suggested the method of cross ruling that was used by Rowland, Wood, and others and probably is the simplest method for obtaining quantitative results. Allen[38] suggested direct observation of the grating in monochromatic light, whereupon periodic errors show up as alternate light and dark bands on the grating face. This method was elaborated by Wood,[26] and appears worth further investigation as a rapid and direct method.†† Michelson[39] measured the periodic error of the screw directly with an automatic method of fringe counting, using an interferometer having one plate on the grating carriage and the other on ways parallel to this. Using this method Gale and Pearson took 120 photographs of the position of the fringes on a cross hair for each turn of the screw. Though very direct and precise this method is slower and more cumbersome than the others. Gale[40] also made use of interference by letting the grating serve as one of the backing mirrors in a Twyman-Green interferometer, while a grating with little periodic error served as the other. Finally, periodic errors can be determined by applying the measured intensities of the Rowland ghosts to their Fourier analysis, as outlined by Meyer[41] on the basis of previous discussions by Rayleigh, Rowland, Anderson, and others.

The method most widely used is that of cross ruling. After some 15 mm of width has been ruled on a small blank, this is rotated through a controlled horizontal angle of the order of 1 min. and returned a distance beyond the position of the first groove equal to between $\frac{1}{4}$ and $\frac{1}{2}$ of the pitch of the screw. A second set of grooves is then ruled over the first. Periodic errors produce a moiré pattern like that shown in Fig. 9, having an amplitude that is a function of the magnitude of the error and the angle between the two sets of grooves. Rayleigh originated this method by putting two small transparent gratings together at a slight angle; if a replica material could be found that would not shrink or expand on curing, it should be possible to test a grating with its own replica. This method of crossed rulings can also be used to test groove straightness.†††

Error of run can be measured by direct counting of fringes, by the multiple-mirror method of Michelson,[39] or by a modified Foucault knife-edge test, as described by Gale.[40] When fringe tests are of long duration it should be remembered that 0.01-in. change in barometric height corresponds to about 1 part in 10^7 in fringe separation.

The groove form produced by a given diamond will not correspond closely to the shape of the ruling edge because of plastic flow. As Babcock[10] and others have shown, direct photo-micrography of grooves is likely to be misleading owing to diffraction effects, though photographs of a single groove

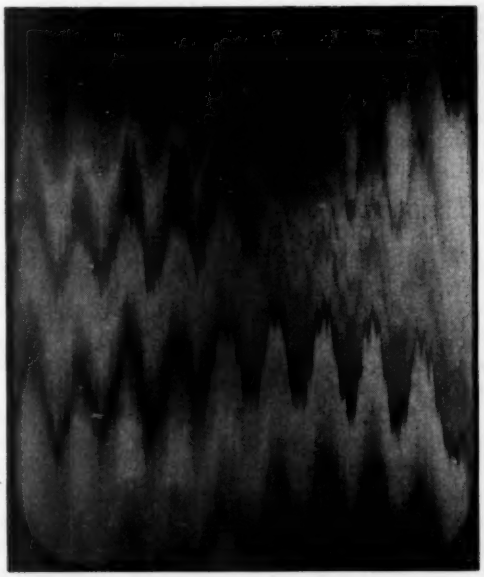

FIG. 9. Typical moiré pattern produced by cross-ruling a small area with an engine having a multiply-periodic error such as to produce Rowland ghosts having intensities about 10 percent of their parent lines in the first order. The amplitude of the periodic error is about 5 micro-inches.

[36] George R. Harrison, See reference 25, p. 58. Harrison, Lord, and Loufbourow, *Practical Spectroscopy* (Prentice-Hall, Inc., New York, 1948), p. 98.

[37] Lord Rayleigh, see reference 17, p. 210 *et seq.*

[38] H. S. Allen, Phil. Mag. **3**, 92 (1902); *ibid.* **6**, 359 (1903).

†† T. J. O'Donnell, on reading this portion of the manuscript, informs me that he has tested plane gratings by illuminating the grating with collimated light from a divider plate, reflecting light back onto the grating with a double prism, and observing the resulting shadow graph through a pinhole in front of the divider plate. Light reflected from the grating at the top edge of one groove is made to strike the bottom of another, displaced two or more grooves to one side. O'Donnell also devised a tilted-slit knife-edge test for concave gratings, and the modified Twyman-Michelson interferometer for testing flat gratings.

[39] A. A. Michelson, J. Frank. Inst. **181**, 785 (1916).

[40] H. G. Gale, Astrophys. J. **86**, 437 (1937).

[41] C. F. Meyer, *The Diffraction of Light, X-rays, and Material Particles* (University of Chicago Press, Chicago, 1934).

††† *Note added in proof:* Babcock points out that a ruling on bare glass, cut at right angles to the grooves and with the two halves superposed in the method of Rayleigh, gives a test free of effects of the first ruling on the second.

crossing the regular grooves at right angles give useful cross-sectional information. Electron micrographs of replicas give an excellent idea of groove form. Probably the most useful instrument for inspection of groove shape is a micro-interferometer of the type made by Bausch & Lomb[41a] or Zeiss.[41b]

CONCLUSION

References to relevant papers that were not cited in the text are given in the appended bibliography. Though recent progress in the improvement of plane and concave replicas has been great,[42] the replica grating is not likely in the near future to eliminate the need for all but original master gratings. Conditions appear to be ripe for a rapid development of automatic engine control methods by means of interference fringes and servo-mechanisms, and these will be discussed in a later paper. Certain conclusions emerge from the considerations of the present paper which point strongly to the desirability and possibility of developing a very coarse plane grating somewhat like á fine reflection echelon with many steps, to be used in orders near the 1000th. The properties of such gratings will be discussed in the next paper of this series.

I am grateful to Messrs. J. A. Anderson, H. D. Babcock, G. H. Briggs, G. H. Dieke, Robert Finn, Edward Leibhardt, George Monk, T. J. O'Donnell, Fred Pearson, Julius Pearson, David Richardson, Manne Siegbahn, John Strong, and R. W. Wood for correspondence regarding unpublished details of early history or the modern accomplishments of ruling engines, and to David Mann, Henry Kurtz, and David Richardson for photographs.

REFERENCES ON GRATING RULING BEYOND THOSE CITED IN THE TEXT

43. Lord Rayleigh, Phil. Mag. **47**, 81, and 193 (1874); *ibid.* **11**, 196 (1881). Experiments with and theory of diffraction gratings.
44. A. Mallock, *Report of Meeting of British Association for Advancement of Science*, p. 466 (1882). Ruling engine for 6″ × 6″ gratings described. See also: Am. Machinist **18**, 241 (1895); *ibid.* **41**, 813 (1914); J. E. Watkins, Smithson. Inst. Report, p. 721 (1890); K. Scheel, Zeits. f. Instr'knde **16**, 321 (1916); W. A. Rogers, Am. J. Sci. **19**, 54 (1880).
45. K. Stöckl, Beiblatt zur Zeits. für Instr'knde **22**, 217 (1903). Suggested improvements in engines.
46. A. S. King, Astrophys. J. **17**, 239 (1903). Ruling errors.
47. A. B. Porter, Phil. Mag. **11**, 154 (1906). Refers to rulings by Max Levy, founder of a firm which has long produced half-tone plates by ruling.
48. A. A. Michelson, Astrophys. J. **18**, 278 (1903). Grating errors.
49. Lord Rayleigh, Proc. Roy. Soc. **A79**, 399 (1907). Dynamical theory of gratings.
50. R. W. Wood, Astrophys. J. **21**, 173 (1905); Physik. Zeits. **6**, 238 (1905). Intensity of spectra.
51. A. Trowbridge and I. B. Crandall, Phil. Mag. **22**, 534 (1911). Experimental study of groove shape in echelettes.
52. F. Biske, Physik. Zeits. **15**, 327 (1914). Theory of groove form.
53. A. Trowbridge, Phys. Rev. **8**, 485 (1916). Echelettes.
54. W. M. Holmes, Industrial Australian and Mining Std., Melbourne **64**, 63 (1920). Testing and ruling of gratings.
55. L. C. Glaser, Zeits. f. Tech. Physik. **7**, 31 (1926). Comparison of fourteen gratings.
56. R. W. Wood and T. Lyman, Phil. Mag. **2**, 310 (1926). Gratings in extreme ultraviolet.
57. E. Buchwald, Ann. d. Physik. **80**, 279 (1926). Ghosts arising from aperiodic errors.
58. J. E. Rosenthal, J. Opt. Soc. Am. **20**, 87 (1930). Theoretical treatment of groove form.
59. C. H. Cartwright, J. Opt. Soc. Am. **21**, 785 (1931). Experiments with laminary grating production.
60. P. Lacroute, Comptes Rendus **194**, 1803 (1932). Study of ghosts.
61. R. W. Ditchburn, Proc. Roy. Irish Acad. **44A**, 123 (1938). Effects of omitting certain grooves.
62. R. C. Williams, Nature **141**, 329 (1938). Ruling in metallic films.
63. C. S. Beals and A. McKellar, J. Roy. Astron. Soc. Canada **32**, 369 (1938). Use of aluminum-on-glass gratings.
64. R. W. Wood, Nature **140**, 723 (1937); *ibid.* **147**, 750 (1941); Popular Astronomy **50**, 146 (1942). Gratings and replicas for astronomy.
65. R. F. Stamm and J. J. Whalen, J. Opt. Soc. Am. **36**, 2 (1946). Calculation of energy distribution with Rowland's formulas.
66. G. K. T. Conn, Proc. Camb. Phil. Soc. **43**, 240 (1947). Theory of the echelette.
67. John Strong, J. Opt. Soc. Am. **39**, 320 (1949). Improvement of echelettes.

[41a] D. Richardson and R. B. Tackaberry, J. Opt. Soc. Am. **38**, 662 (1948).
[41b] W. Kinder, Zeiss. Nachtrichten **3**, 91 (1937).
[42] R. W. Wood, J. Opt. Soc. Am. **36**, 715 (1946).

A Comparison of Electrical and Psychophysical Determinations of the Spectral Sensitivity of the Human Eye

LORRIN A. RIGGS, RICHARD N. BERRY, AND MATTHEW WAYNER*
Psychological Laboratory, Brown University, Providence, Rhode Island
(Received February 28, 1949)

Electrical responses of the human retina have been measured by the use of an electrode mounted on a contact lens. The form and magnitude of such responses have been determined for stimulation by filtered lights having various dominant wave-lengths. Two sets of spectral sensitivity data, one for the dark-adapted eye and the other for the light-adapted, have been obtained by computing the intensity of stimulation necessary at each wave-length to arouse an electrical response of a given small magnitude. Comparable data have been obtained in psychophysical experiments in which the same filter combinations and closely similar methods of computation have been employed. The electrical data, for both the light-adapted and the dark-adapted eye, agree much more closely with the psychophysically determined scotopic sensitivity curve than with the photopic. It appears, however, that lights of the shorter wave-lengths are somewhat more effective in arousing electrical responses than the scotopic sensitivity curve would predict.

INTRODUCTION

SUCH visual functions as dark adaptation, light adaptation, and spectral sensitivity have traditionally been explored by psychophysical methods. Recently, by means of an objective technique involving an electrode imbedded in a contact lens,[1] it has become possible to record electrical responses originating in the retina itself. The present experiments afford an opportunity to compare these two very different methods of approach to a single visual problem. The problem selected for this comparison is that of determining the sensitivity of the human eye to selectively filtered lights having various dominant wave-lengths within the visual region of the spectrum.

The spectral sensitivity of the eye has been expressed traditionally in the form of two curves which relate sensitivity to wave-length. One of these curves, the scotopic, concerns the rod receptors as tested by stimuli at very low levels of brightness in the dark-adapted eye. The other, the photopic, has been variously determined by photometric measurements of the cone receptors, usually within the fovea. A "standard" photopic curve, adopted in 1924 by the International Commission on Illumination, has been widely employed[2] for specifications of photopic luminosity. It has been used in this way as a preliminary basis for calculating the photopic results reported below. A "mean scotopic visibility curve," based upon five determinations by separate investigators, has been proposed as standard for this function by Stiles and Smith.[3] This curve has similarly served as a basis for the scotopic calculations of the present paper.

The most recent series of psychophysical determinations of human spectral sensitivity has been reported by Wald.[4] As a result of experiments on the absolute threshold of the foveal cones, Wald concluded that their point of maximum sensitivity lies at about 562 mμ. His results for this function, in common with some of the best of the earlier determinations, depart rather widely from the I.C.I. curve, particularly at the shorter wave-lengths. Wald has also determined the spectral sensitivity function of extra-foveal cones by working during the interval of dark adaptation when the cones determine the threshold. Under these conditions, the maximum sensitivity is at 555 mμ and the curve is relatively higher than that of the fovea between 440 and 490 mμ. Wald points out that the discrepancy represents the absorption of these rays by the macular pigment. Wald's curve for the peripheral rods shows a maximum sensitivity at about 505 mμ and does not differ greatly from those of other investigators who have determined absolute thresholds or low level brightness matches for the dark-adapted eye.

In the experiments reported below both electrical and psychophysical data were secured on five observers using the same set of filtered light stimuli. Four specific methods were employed, as follows: (1) Electrical responses of the retina were measured in the fully dark-adapted eye in order to find the relative energy of stimulation, at various wave-lengths, which would arouse a constant magnitude of response. (2) Similar electrical responses were obtained for the light-adapted retina. (3) A scotopic sensitivity function was determined by a conventional low brightness matching procedure in the fully dark-adapted eye. (4) A photopic sensitivity function was determined by flicker pho-

* This research was carried out under Contract N7onr-358, Task Order II, between the ONR and Brown University.
[1] L. A. Riggs, Proc. Soc. Exper. Biol. Med. **48**, 204 (1941).
[2] *I.E.S. Lighting Handbook* (Illuminating Engineering Society, New York, 1947).
[3] W. S. Stiles and T. Smith, Proc. Phys. Soc. Lond. **56**, 251 (1944).
[4] G. Wald, Science **101**, 653 (1945).

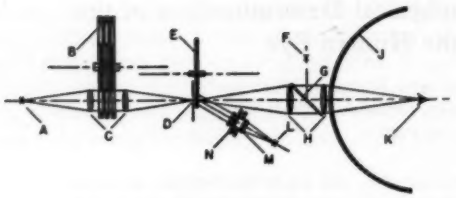

FIG. 1. A schematic drawing (not to scale) of the apparatus used to provide visual stimulation. A—tungsten filament source; B—filters in rotary mounts; C—first pair of collimating lenses; D—diaphragm; E—rotary shutter or flicker vane; F—fixation point; G—glass plane for viewing fixation point; H—second pair of collimating lenses; J—bowl-shaped adapting screen; K—the eye; L—tungsten filament source for illuminating flicker vane; M—collimating lens; N—variable density wedges.

tometry at a high level of brightness in the light-adapted eye.

APPARATUS

Optical System

The principal function of the optical system (Fig. 1) is to provide the eye with flashes of selectively filtered light of moderate to high intensity. The light source is a 6-v, 50-C.P. tungsten-filament lamp, type 1183, run from a storage battery at a filament temperature of 3300°K. The selective filters are used in a collimated portion of the beam from this source. They are mounted on disks which may be rotated to bring any desired filter into the beam. The filters are shielded from the light by the use of an opaque vane which is deflected out of the beam momentarily before the stimulus flash is to be delivered. An image of the filament is formed in the plane of a small diaphragm. Here, a circular aperture serves to limit the effective area of the filament to a disk, 1.5 mm in diameter, including only the hottest central portions of its two coils. The light is intercepted in a plane close to this aperture by the use of a gravity-operated disk shutter which provides a flash having a duration of 0.04 sec. Alternatively, this disk is replaced by a white flicker vane providing two flashes per revolution. The rate of rotation of the flicker vane is under the control of the experimenter and is observed by the use of a tachometer. An auxiliary source, also a battery-operated type 1183 lamp, provides 3300° white illumination on the flicker vane for the photometric measurements at high levels of brightness.

An image of the filament coils and the occluding diaphragm is formed in the plane of the pupil by the use of two additional achromatic lenses. Between the latter is inserted a plane of thin glass at 45° for the purpose of providing the observer with a fixation point. Fixation is on a small, dim red filament appearing at the center of the stimulus patch. The patch is seen as a Maxwellian view of the last lens, having a diameter of 7°30′. When desired, a diffuse surrounding field is provided by illuminating a bowl-shaped, white-painted surface by the use of lamps situated near the observer's head. The eye is carefully adjusted to the correct position by the use of a biting-board and head rest. Thus all the rays emerging from the optical system pass through the center of the pupil.

Recording System

The manner of recording retinal action potentials has been described elsewhere.[5] Briefly, it involves the use of a contact lens to support an electrode. A 3-mm hole is drilled through the outer nasal portion of the corneal bulge on a plastic contact lens. A silver disk is fitted into this hole in such a way as to make contact with the fluid between the contact lens and the cornea. A light, flexible wire connects this electrode with the input grid of a high gain, direct-coupled amplifier. The input ground is applied to a point on the forehead. The output of the amplifier is used to drive a loop oscillograph galvanometer, photographic traces from which constitute the experimental records of the electrical responses of the retina.

CALCULATIONS OF LUMINOUS FLUX

The spectral sensitivity of the eye, as Griffin, Hubbard, and Wald[6] have pointed out, is defined as the reciprocal of the relative energy required at each wave-length to elicit a constant visual effect. Using their notation, the luminous flux provided to the observer by a selectively filtered beam is given by $\int R_\lambda K_\lambda t_\lambda d\lambda$. In this expression, R_λ represents the energy emitted by the source at any value, λ of wave-length; t_λ is simply the transmission of the filter or filters at any value of λ; and K_λ is the spectral sensitivity of the eye at any value of λ. The calculations appearing below are based upon the following data for the three significant variables:

The Source Factor, R_λ.—The apparent black-body temperature of the tungsten filament was measured as 2908°K by the use of a Leeds and Northrup Optical Pyrometer.[7] Interpolation from a table published by Harnwell and Livingood[8] yields a value of about 3300°K as the true temperature of the tungsten filament. From the International Critical Tables are obtained the necessary data on

[5] See Reference 1 above.

[6] D. R. Griffin, R. Hubbard, and G. Wald, J. Opt. Soc. Am. **37**, 546 (1947).

[7] We are indebted to Professor H. E. Farnsworth of the Department of Physics, Brown University, for lending us the pyrometer used in these experiments.

[8] G. P. Harnwell and J. J. Livingood, *Experimental Atomic Physics* (McGraw-Hill Book Company, Inc., New York, 1933), p. 206.

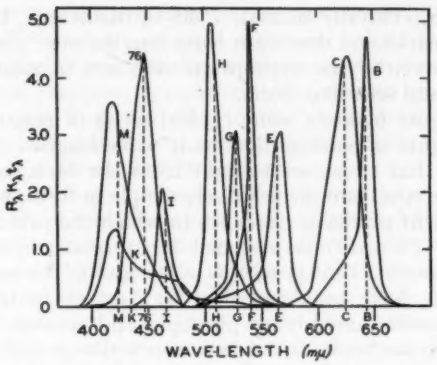

FIG. 2. Calculated values of $R_\lambda K_\lambda t_\lambda$, the scotopic visual effectiveness of the radiation transmitted by the selective filters used in this investigation. Filter 76 is the violet filter of the Wratten monochromatic series. Filter *M* is composed of a Corning glass filter plus an interference type filter supplied by the Farrand Optical Company. The other filters, which we have arbitrarily designated by capital letters, are Farrand interference type filters. The dotted line bisects the area under each curve. It denotes the "center of gravity" wave-length for each filter. In drawing each of the curves the value of $R_\lambda K_\lambda t_\lambda$ has been multiplied by a factor as follows: *M*—2000; *K*—400; 76—1000; *I*—40; *H*—40; *G*—20; *F*—40; *E*—40; *C*—1000; *B*—4000.

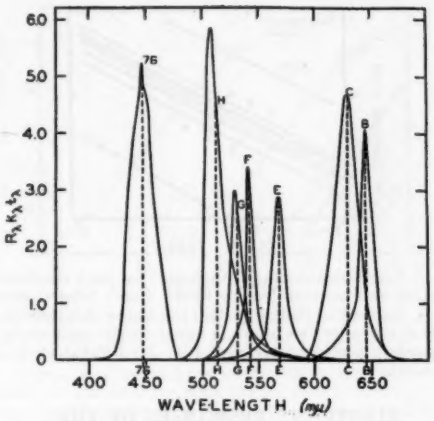

FIG. 3. Calculated values of $R_\lambda k_\lambda t_\lambda$, the photopic visual effectiveness of the radiation transmitted by the filters used in this investigation. Each value has been multiplied by a factor as follows: 76—10,000; *H*—100; *G*—20; *F*—20; *E*—10; *C*—40; *B*—100.

the emissivity of tungsten at that temperature at each value of λ. This factor is then multiplied by the ideal black-body emission for each wave-length to give the true emission of a 3300° tungsten source. For the sake of convenience, R_λ is set equal to unity when $\lambda = 700$ mμ; all other values of R_λ are thus relative to the value at that wave-length.

The Filter Transmission Factor, t_λ.—All the filters used in this investigation were calibrated by the use of a Beckman DU Spectrophotometer.[9] It was possible to use very narrow wave-length bands in the regions of peak transmission, so that accurate data were secured on even the extremely narrow peaks of the Farrand filters. For the regions of low transmission wider bands were used, and transmissions as low as 0.0005 were measurable.

The Spectral Sensitivity Factor, K_λ.—For scotopic calculations the relative values of Stiles and Smith[10] were used. The maximum value of *K*, occurring at $\lambda = 510$ mμ, was set at unity. In Fig. 2 are plotted the values of $R_\lambda K_\lambda t_\lambda$ for each of the selective filters used in the scotopic determinations. The total luminous flux in each case is simply the area under the curve, i.e., $\int R_\lambda K_\lambda t_\lambda d\lambda$. For photopic calculations the I.C.I. photopic luminosity factors were employed. These we shall call k_λ, to distinguish them from their scotopic counterparts, K_λ. The corresponding curve for each filter is shown in Fig. 3. For each of the curves in Figs. 2 and 3 a "center of gravity" has been defined as the wave-length above or below which is found one-half of the area under the curve.

The necessary variation in intensity of stimulation was achieved by the insertion of Wratten "neutral tint" filters into the beam. On the basis of spectrophotometric measurement of these filters, it was possible to construct curves similar to those of Figs. 2 and 3 for various combinations of selective and "neutral" filters, in order to evaluate the luminous flux for each of the combinations. Two master charts were thus developed, as shown in Figs. 4 and 5, to show the relation of luminous flux to rated filter density. The slightly differing slopes of line for the various filters are an indication of the fact that the "neutral" filters are themselves partially selective in their spectral transmission.

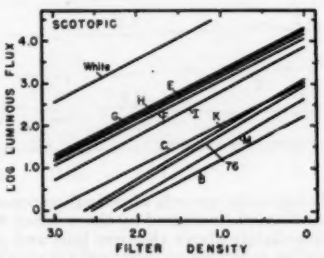

FIG. 4. Log luminous flux (scotopic) for each combination of selective and neutral-density filters. For a filter density of zero (i.e., no neutral-density filter) the value of luminous flux for any of the selective filters is equal to the area under the corresponding scotopic curve of Fig. 2 ($\int R_\lambda K_\lambda t_\lambda d\lambda$ in arbitrary units).

[9] It may be noted here that some erroneous filter calibrations were made earlier on a G. E. recording spectrophotometer. This instrument, as Middleton has recently pointed out [J. Opt. Soc. Am. **38**, 74 (1948)], is unsuited unless specially modified for the measurement of transmission of samples which have a high specular reflectance.

[10] See Reference 3 above.

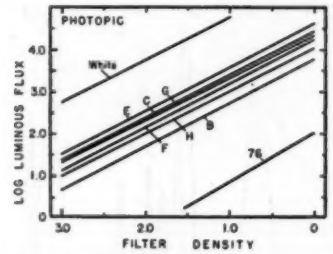

FIG. 5. Log luminous flux (photopic) for each combination of selective and neutral-density filters. For a filter density of zero (i.e., no neutral-density filter) the value of luminous flux for any of the selective filters is equal to the area under the corresponding photopic curve of Fig. 3 ($\int R_\lambda k_\lambda l_\lambda d\lambda$ in arbitrary units).

ELECTRICAL RESPONSES OF THE DARK-ADAPTED EYE

Electrical responses to stimulation by light were recorded from the right eye of each subject by the method outlined in an earlier section of this paper. The eye was carefully centered with respect to the stimulating beam by the adjustment of the biting board, and the observer was allowed to dark adapt for a minimum of 45 minutes. The spectral sensitivity was then found by measuring the magnitude of the electrical response to each of a series of test flashes in the dark-adapted eye. The flashes were of uniform duration (0.04 sec.) and area (circular patch, centrally fixated, 7°30′ in diameter). Their intensities and dominant wave-lengths were varied, however, by the appropriate selection of neutral-tint and selective filters.

Figure 6 shows some typical series of responses to lights of various dominant wave-lengths. It is seen that responses to light from the blue region of the spectrum are relatively simple in form, while red light produces responses in which the principal wave (*B*-wave) is preceded by a brief, diphasic component. Effects of this sort, specific to wave-length, have been described by Adrian[11] in terms of possible underlying photopic and scotopic receptor mechanisms. Adrian's conclusion is that the first peak of this double-peaked *B*-wave is photopic in origin, while the second peak, which is ordinarily the dominant one, is scotopic. Only red light favors the photopic system sufficiently to allow the small photopic component to escape domination by the scotopic.

For a quantitative indication of the magnitude of the electrical response we have chosen to measure the height of the *B*-wave. Figure 7 shows a series of measurements of *B*-wave heights as influenced by variations in the intensity of stimulation. It may be seen from this figure that these variations have been achieved by the insertion of Wratten neutral-tint filters having density ratings of 0.3, 0.6, 1.0, 1.3, 1.6, 2.0, etc. log units. For

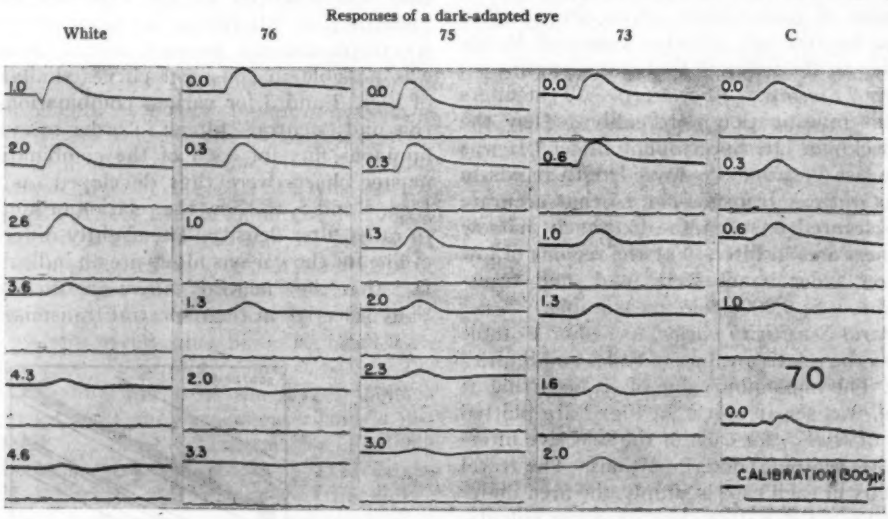

FIG. 6. Photographic records of retinal responses to illumination by white and colored lights of various intensities. In each record time (0.1 sec.) is indicated by the notched horizontal line, the moment of the flash is denoted by the short horizontal line appearing immediately above the time line, and the electrical response is indicated by the oscillographic trace, in which a rise above the baseline indicates an increased positivity of the cornea with respect to the fundus. Filters 76, 75, 73, and 70 are Wratten filters which selectively transmit violet, blue-green, yellow-green, and deep-red light, respectively. Filter *C* is a Farrand interference type filter transmitting orange-red light. The number appearing over each trace denotes the rated density (logarithmic units) of the Wratten neutral-tint filter employed. The calibration indicates the effect of applying a 300-microvolt square wave to the input of the direct-coupled amplifier. Subject FR; experiments of July 20 and July 22, 1948.

[11] E. D. Adrian, J. Physiol. **104**, 84 (1945), and J. Physiol. **105**, 24 (1946).

reasons which have been enumerated in previous papers[12,13] we have adopted a *B*-wave height of 58 microvolts as representing a minimum measurable electrical response. A horizontal line may be drawn across Fig. 7 at the 58-microvolt level, and the corresponding values of neutral filter density may be read off for each of the curves.

Curves similar to those of Fig. 7 were later obtained for all observers by the use of selected filters of the Farrand interference type. Certain Corning glass and Wratten gelatin filters of the "monochromatic" series were also employed. The most highly selective of all these filters were chosen for the following electrical determinations of the spectral sensitivity of the dark-adapted retina.

Table I shows the values of log luminous flux required to elicit the standard 58-microvolt response in the dark-adapted retina. Each of these values was read off from Fig. 4 after the necessary combination of filters to elicit this response was determined from an experimental curve similar to those of Fig. 7. For an example, a neutral tint filter density of 2.31 log units in conjunction with the *H* filter produced the 58-microvolt response in the retina of observer LSH. Figure 4 shows that for this filter combination log luminous flux = 1.92, and this is the value given in Table I.

Also shown in Table I is the departure of each of the log luminous flux values from the one for the *H* filter at 509 mμ. It is obvious that these departures would all be close to zero if the electrical data were in good agreement with the Stiles and Smith scotopic data on relative spectral sensitivity. Instead, the electrical data reveal a definitely higher sensitivity (i.e., lower values of luminous flux) in the blue region of the spectrum, and a somewhat lower sensitivity in the red.

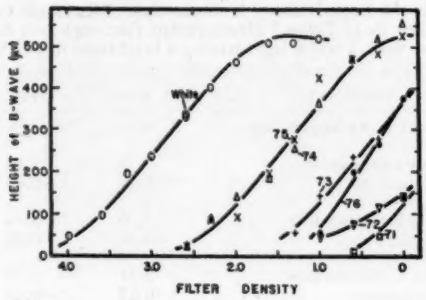

FIG. 7. Height of principal component (*B*-wave) in retinal responses to white light and to the colored lights of the Wratten monochromatic series. The density of the Wratten neutral density filters is given in logarithmic units. The curves are all parallel to one another with the exception of 76 and 72. The curve for 76 is steeper because the neutral filters absorb blue light more than light of other wave-lengths. The shape of 72 is complicated by the presence of the double-peaked *B*-wave. Subject LSH; experiments of May 7 and May 14, 1948.

Figure 8 presents a more direct comparison between the electrical data and the scotopic sensitivity curve of Stiles and Smith. Also shown in this figure is the sensitivity curve of Wald[14] for the peripheral rods. Each of the experimental points in this figure is plotted as a departure from the Stiles

TABLE I. Logarithms of luminous flux, $\int R_\lambda t_\lambda K_\lambda d\lambda$ to elicit electrical responses of 58 microvolts in the dark-adapted retina. Test flashes of 0.04-sec. duration, central fixation on a circular field, 7°30′ in diameter. Contact lens electrode over cornea, neutral electrode over supra-orbital ridge. Experiments of July, 1948 and December, 1948 (the latter data are marked with asterisks).

Selective filter	*M*	*K*	76	*I*	*H*	*G*	*F*	*E*	*C*
Dominant wave-length (mμ)	424	434	446	466	509	529	541	564	625
LSH log luminous flux	1.30*	1.50	1.51	1.85	1.92	2.02	2.04	2.12	2.14
departure	−0.62	−0.42	−0.41	−0.07	0.00	0.10	0.12	0.20	0.22
LML log luminous flux	1.10*	1.14	1.12	1.35	1.53	1.62	1.93	2.12	2.03
departure	−0.77	−0.39	−0.41	−0.18	0.00	0.09	0.40	0.59	0.50
FR log luminous flux	0.84*	1.00	1.15*	1.35	1.53	1.37	1.58	1.87	1.82
departure	−0.69	−0.58	−0.38	−0.18	0.00	−0.16	0.05	0.34	0.29
LAR log luminous flux		1.20	1.17	1.58	1.73	1.83	1.73	2.14	2.43
departure		−0.53	−0.56	−0.15	0.00	0.10	0.00	0.41	0.70
MJW log luminous flux		1.35		1.35	1.79	1.96	1.81	2.30	2.32
departure		−0.44		−0.44	0.00	0.17	0.02	0.51	0.53
Mean departures (from the Stiles and Smith scotopic curve relative to the value at 509 mμ).	−0.69	−0.46	−0.44	−0.20	0.00	0.06	0.12	0.41	0.45

[12] E. P. Johnson, J. Exper. Psych., in press (1949).
[13] L. A. Riggs and E. P. Johnson, J. Exper. Psych., in press (1949).
[14] These curves have been redrawn from those appearing in the original articles, References 4 and 3 above, respectively. For purposes of comparison a sensitivity of unity has been assigned to all values at 509 mμ.

TABLE II. Logarithms of luminous flux, $\int R_\lambda t_\lambda K_\lambda d\lambda$ to elicit electrical responses of 58 microvolts in the light-adapted retina. Conditions as in Table I above except that each test flash was delivered 60 seconds after a two-minute exposure to a large adapting field of white light having a brightness of 202 footlamberts. Experiments of August, 1948.

Selective filter	*K*	76	*I*	*H*	*G*	*F*	*E*	*C*
Dominant wave-length (mμ)	434	446	466	509	529	541	564	625
RNB log luminous flux	2.74	2.66	2.86	3.14	3.35	3.38	3.30	3.00
departure	−0.40	−0.48	−0.28	0.00	0.21	0.24	0.16	−0.14
LSH log luminous flux	3.10	3.31	3.02	3.30	3.40	3.52	3.76	3.15
departure	−0.20	0.01	−0.28	0.00	0.10	0.22	0.46	−0.15
LML log luminous flux	2.21	2.25	2.55	2.74	3.10	3.01	3.51	2.59
departure	−0.53	−0.49	−0.19	0.00	0.36	0.27	0.77	−0.15
FR log luminous flux	2.64	2.44	3.13	3.04	3.25	3.22	3.47	2.77
departure	−0.40	−0.60	0.08	0.00	0.21	0.18	0.43	−0.27
LAR log luminous flux	3.66		3.55	3.75	3.90	3.72	4.08	3.45
departure	−0.09		−0.20	0.00	0.15	−0.03	0.33	−0.30
Mean departures (from the Stiles and Smith scotopic curve relative to the value of 509 mμ)	−0.32	−0.39	−0.17	0.00	0.21	0.18	0.43	−0.20

and Smith curve, as given by Table I. For an example, the departure of the electrical data of LSH from the Stiles and Smith curve is given as −0.62 for the *M* filter. This point is at a vertical distance of 0.62 log unit from the Stiles and Smith curve in the direction of higher sensitivity. It is clear that while there is not a perfect agreement among the data for the five observers there is a consistent tendency for the points to depart from the Stiles and Smith curve in the direction of higher sensitivity in the blue region of the spectrum.

ELECTRICAL RESPONSES OF THE LIGHT-ADAPTED RETINA

It was originally intended to record the electrical responses of the foveal region of the retina in this section of the experiments. Data so obtained were to have been compared with the psychophysical data appearing below in the section on photopic determinations by the method of flicker photometry. Adrian[11] has noted, however, that no electrical responses are observed when a moderate level of stimulation is applied to the fovea. With very high levels of stimulation, it is probable that such responses as are obtainable may largely be the result of stimulating the para-foveal portions of the retina by scattered light. Furthermore, it has been found in previous experiments[13] conducted in this laboratory that even a relatively large region of the retina fails to respond during or immediately following steady adaptation to a field of moderate intensity. These considerations have made it necessary to employ a relatively large area of stimulation, and to allow a brief interval of dark adaptation to follow exposure to the adapting light.

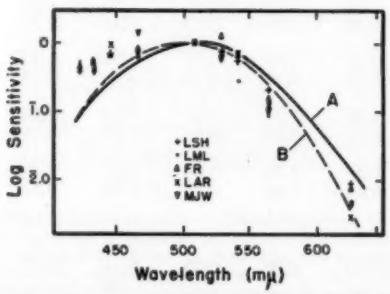

FIG. 8. Log sensitivity as indicated by electrical responses of the dark-adapted eye. Data plotted from Table I by the method described in the text. Curve *A* is the mean scotopic curve of Stiles and Smith. Curve *B* is the curve of Wald for the peripheral rods.

To meet the above-mentioned requirements, the eye was first dark adapted and then exposed for two minutes to an adapting field of white light having a brightness of 202 footlamberts. After exactly one minute of dark adaptation following this two-minute interval in the light, a test flash from the main optical system was delivered to the eye. This flash had the same characteristics as the ones used for the dark-adapted eye: A diameter of 7°30′ visual angle, a duration of 0.04 sec., intensity and dominant wave-length dependent upon the choice of filters. Again the eye was turned to the light-adapting field for two minutes, and again a test flash was delivered after one minute of dark adaptation. In this way it was possible to obtain a series of records from the eye at a relatively constant level of light adaptation. It may be noted that the general procedure here was similar to the one used by Wald[15] in measuring the spectral sensitivities of the peripheral cones.

Figure 9 shows some typical responses of the

[15] See Reference 4 above.

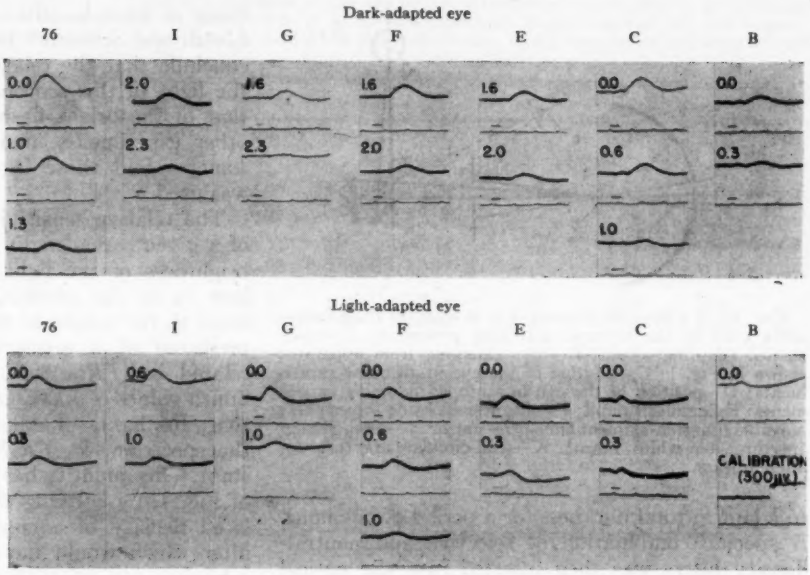

FIG. 9. Comparative responses of the dark-adapted and light-adapted eye. Subject LML; experiments of June 17, August 17, August 19, and August 23, 1948.

light-adapted eye as compared with similar records from the dark-adapted eye. It is seen that light adaptation has served to reduce the magnitude of the *B*-wave. At the same time the latency of the response has clearly been reduced, and some changes in wave form are apparent. Table II presents the values of log luminous flux, assuming again the Stiles and Smith scotopic factors, which would have been required to elicit the standard 58-microvolt response from the light-adapted eye of each observer. These values have been derived from curves, for the light-adapted eye, similar to those of the dark-adapted eye appearing in Fig. 7. Figure 10 shows again, for the values so obtained, the extent of their agreement with the scotopic curves of Wald and of Stiles and Smith. The data for Fig. 10 have been taken from Table II.

It might have been anticipated that the light-adapted eye would exhibit a spectral sensitivity curve which would differ considerably from its dark-adapted counterpart. The Purkinje shift fails to appear, however, and it is seen that in the light-adapted eye the electrical data appear to resemble those of the dark-adapted eye in showing an unexpectedly high sensitivity in the blue region of the spectrum. At only one point, namely that of the *C* filter in the red-orange region, is there an indication of the expected shift toward greater relative sensitivity at the longer wave-lengths. In this case, however, measurements of the height of the *B*-wave are complicated by the presence of the double peak which is characteristic of responses to red light, as may be seen from the records of Figs. 6 and 9.

DETERMINATIONS OF SCOTOPIC SENSITIVITY

The psychophysical method of adjustment was used for measuring the spectral sensitivity of the eye at a low level of brightness. In these experiments the eye was not placed in the position shown in Fig. 1. Instead, some of the light emerging from the optical system was allowed to fall as a semicircular patch on a disk coated with a film of MgO. This arrangement is shown in Fig. 11. Also shown in this drawing is an arrangement by which the other half of the disk was illuminated by white light from a movable source consisting of a calibrated lamp whose distance from the screen was under the control of the observer. The task of the observer was to equate the brightnesses of the two halves of the field by adjusting this distance.

The scotopic effectiveness of the selectively filtered light was determined for each observer by the following procedure: The eye was first allowed

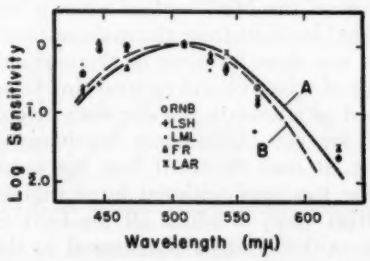

FIG. 10. Log sensitivity as indicated by electrical responses of the light-adapted eye. Data plotted from Table II by the method described in the text. Curve *A* is the mean scotopic curve of Stiles and Smith. Curve *B* is the curve of Wald for the peripheral rods.

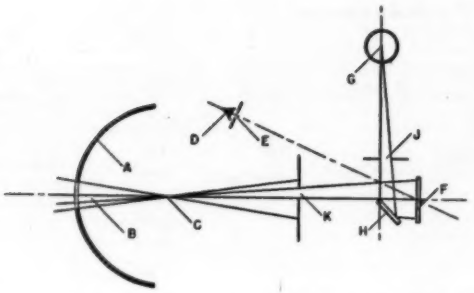

FIG. 11. A schematic drawing (not to scale) of the arrangement used in the scotopic matching procedure. A—bowl-shaped adapting screen; B—rays emerging from apparatus shown in Fig. 1; C—position of the eye in all other experiments; D—position of the eye in scotopic matching experiments; E—artificial pupil; F—magnesium oxide screen; G—movable tungsten-filament lamp; H—mirror; J—semi-circular aperture for white beam; K—semi-circular aperture for filtered beam.

to adapt in total darkness for a period of one hour. A specific combination of selective and neutral-density filters was then inserted into the optical system, and the resulting light was allowed to fall on one-half of the MgO disk shown at F in Fig. 11. The observer was then instructed to view the disk through an artificial pupil at E and to move the white comparison lamp, G, until the two halves of F appeared to match one another in brightness. The optical distance of G from F was then read by the experimenter, and the brightness for this distance was computed from the known candlepower of the lamp and reflectance of the surface. Each observer made three settings on each of nine filter combinations. The brightnesses so measured for the selective filters with appropriate neutral-density filters ranged between 10 and 30 microfootlamberts, a level below the photo-chromatic interval for all except the deep red light. Lights of other wave-lengths appeared uniformly achromatic to all observers at this level of brightness, and it was not difficult for the observers to make the required adjustment of equality between the semicircles.

As indicated above, each observer determined the brightness of the MgO surface when it was dimly illuminated by light from the main optical system; and this was done for some of the particular combinations of selective and neutral-tint filters which were used to stimulate the eye with bright flashes of light for the experiments involving electrical recording. It may be noted that the reduction in brightness has been achieved by changing from a Maxwellian view, in which all the light emerging from the optical system is delivered to the eye of the observer, to the viewing of an opaque disk under conditions such that only 1.2×10^{-6} of the original light is so delivered. The reflectance of a MgO surface is substantially constant over the range of wave-lengths employed here, as Benford, Lloyd, and Schwarz[16] have shown. Hence we may conclude that the relative spectral distribution of the light in this scotopic experiment is similar to that of the bright flashes of light used in all the other experiments, namely those on flicker photometry and those in which electrical recording was used.

The relative sensitivity of an observer to light of a given wave-length is obviously related to the brightness of the light as seen by the observer, just as in the electrical experiments it was related to the height of the B-wave. The brightness predicted of a given filter combination is also related to $\int R_\lambda K_\lambda t_\lambda d\lambda$, assuming the Stiles and Smith values of K_λ. Consequently it may be stated that, for lights of equal brightnesses throughout the spectrum, $\log \int R_\lambda K_\lambda t_\lambda d\lambda$ should remain constant. This quantity has been found, from the chart of Fig. 4, by entering this chart with the interpolated density of a continuously variable neutral filter which would have allowed each subject to achieve a standard brightness of 22 microfootlamberts. The latter estimate is obviously possible on the basis of interpolation from the brightnesses determined by each observer for the fixed neutral density filters actually used with each of the selective filters.

Table III presents the log of $\int R_\lambda K_\lambda t_\lambda d\lambda$, the luminous flux required by each of the observers to attain a standard scotopic brightness by the use of an appropriate density of neutral filter. It is seen that the values are not perfectly constant, but tend to be slightly low in the blue region and higher in the red. This is interpreted to mean that in these experiments the sensitivity for blue light is somewhat higher than that predicted by the Stiles and Smith function, while for red light it is lower.

Figure 12 has been plotted from the data of Table III by the same scheme as that used in plotting the electrical data in Fig. 8, from the values given by Table I. It is seen that the individual points of the five observers used here are in relatively good agreement with one another, and with the scotopic curves of Stiles and Smith and of Wald.

DETERMINATIONS OF PHOTOPIC SENSITIVITY

Data on photopic sensitivity were obtained by the use of flicker photometry. The bowl-shaped surrounding field was illuminated by white light to achieve a brightness of 227 footlamberts. This is approximately the same brightness as that of the stimulus flashes which were used in the electrical response experiments. A flicker vane, inserted at E in Fig. 1, was illuminated by white light from the auxiliary source. Variations in the intensity of this

[16] Benford, Lloyd, and Schwarz, J. Opt. Soc. Am. **38**, 445 (1948).

TABLE III. Logarithms of luminous flux, $\int R_\lambda t_\lambda K_\lambda d\lambda$, delivered to the eye by the arrangement shown in Fig. 11. Scotopic matching by five observers of a stimulus patch having a brightness of 22 microfootlamberts. Experiments of July and August, 1948. Dark-adapted eye, central fixation on circular bipartite field 8° in diameter. Values of K_λ from Stiles and Smith mean scotopic data.

Selective filter	*K*	76	*I*	*H*	*G*	*F*	*E*	*C*	*B*
Dominant wave-length (mμ) (see Fig. 2)	434	446	466	509	529	541	564	625	643
RNB log luminous flux	$\bar{5}.79$	$\bar{5}.86$	$\bar{4}.09$	$\bar{4}.10$	$\bar{4}.18$	$\bar{4}.29$	$\bar{4}.07$	$\bar{4}.20$	$\bar{4}.31$
departure	−0.31	−0.24	−0.01	0.00	0.08	0.19	−0.03	0.10	0.21
LSH log luminous flux	$\bar{5}.80$	$\bar{5}.88$	$\bar{5}.93$	$\bar{5}.98$	$\bar{4}.15$	$\bar{4}.10$	$\bar{4}.14$	$\bar{4}.10$	$\bar{4}.30$
departure	−0.18	−0.10	−0.05	0.00	0.17	0.12	0.16	0.12	0.32
LML log luminous flux	$\bar{5}.81$	$\bar{5}.82$	$\bar{5}.97$	$\bar{4}.14$	$\bar{4}.06$	$\bar{4}.02$	$\bar{4}.10$	$\bar{4}.08$	$\bar{4}.19$
departure	−0.33	−0.32	−0.17	0.00	−0.08	−0.12	−0.04	−0.06	0.05
FR log luminous flux	$\bar{5}.70$	$\bar{5}.77$	$\bar{5}.77$	$\bar{5}.92$	$\bar{5}.95$	$\bar{5}.81$	$\bar{5}.93$	$\bar{5}.98$	$\bar{5}.95$
departure	−0.22	−0.15	−0.15	0.00	0.03	−0.11	0.01	0.06	0.03
LAR log luminous flux	$\bar{5}.76$	$\bar{5}.81$	$\bar{5}.79$	$\bar{5}.92$	$\bar{5}.96$	$\bar{5}.85$	$\bar{4}.00$	$\bar{4}.07$	$\bar{4}.05$
departure	−0.16	−0.11	−0.13	0.00	0.04	−0.07	0.08	0.15	0.13
Mean departures (from the Stiles and Smith scotopic curve relative to the value at 509 mμ)	−0.24	−0.18	−0.10	0.00	0.05	0.00	0.04	0.07	0.15

comparison light were achieved by the use of an optical wedge which was under the control of the observer.

As the flicker vane rotated, it presented the observer with alternate flashes of this white comparison light and the selectively filtered light from the main optical system. The brightness of the flickering field was similar to that of the steady surrounding field. The task of the observer was to adjust the wedge to a point where flickering was at a minimum for each of the fixed filter combinations employed in these experiments. Central fixation was employed. The circular flickering field was reduced to 1°50′ in diameter to insure foveal stimulation. An artificial pupil having a diameter slightly larger than that of the image from the main beam but smaller than the natural pupil was placed in front of the eye to eliminate the effects of pupillary contraction in connection with the comparison light.

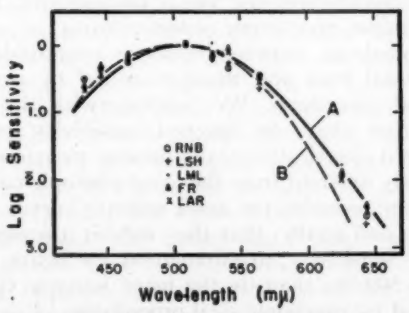

FIG. 12. Log sensitivity measured by a scotopic matching procedure. Data plotted from Table III by the method described in the text. Curve *A* is the mean scotopic curve of Stiles and Smith. Curve *B* is the curve of Wald for the peripheral rods.

As in the case of the scotopic determinations above it was necessary to compute, for each observer, the amount of luminous flux of the selectively filtered light which would have resulted in a certain standard setting of the comparison field. In this case the computation was based upon the wedge setting actually made by each observer for each of the fixed filter combinations. Since the transmission of the wedge was known for each of these settings it was possible to specify the precise amount of luminous flux which would have resulted in a minimum degree of flicker at the standard setting of the wedge. The logarithms of these luminous flux values appear in Table IV.

Figure 13 shows graphically the extent to which the photopic flicker data conform to the photopic spectral sensitivity function of Wald and that of the I.C.I. The data are plotted here by the same general scheme as employed in Fig. 12 for the scotopic data. In Fig. 13, however, the points have been plotted as departures from the I.C.I. curve for photopic luminosity as given in Table IV.

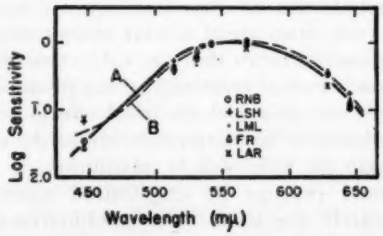

FIG. 13. Log sensitivity measured by flicker photometry at a photopic level. Data plotted from Table IV by the method described in the text. Curve *A* is the I.C.I. photopic luminosity curve. Curve *B* is the curve of Wald for the foveal cones.

TABLE IV. Logarithms of luminous flux, $\int R_\lambda t_\lambda k_\lambda d\lambda$ for photopic matching of a stimulus patch having a brightness of 227 footlamberts. Light-adapted eye, method of flicker photometry with central fixation on a circular flickering field 1°50′ in diameter surrounded by a large adapting field of white light having a brightness of 227 footlamberts. Experiments of August, 1948. Values of k_λ from I.C.I. photopic luminosity data.

Selective filter	76	*H*	*G*	*F*	*E*	*C*	*B*
Dominant wave-length (mμ) (see Fig. 3)	447	513	531	542	568	628	645
RNB log luminous flux	1.68	1.89	1.89	1.71	1.68	1.66	1.86
departure	−0.03	0.18	0.18	0.00	−0.03	−0.05	0.15
LSH log luminous flux	1.50	1.88	2.01	1.82	1.85	1.76	1.75
departure	−0.32	0.06	0.19	0.00	0.03	−0.06	−0.07
LML log luminous flux	1.74	1.96	1.99	1.71	1.76	1.59	1.80
departure	0.03	0.25	0.28	0.00	0.05	−0.12	0.09
FR log luminous flux	1.85	1.90	1.85	1.78	1.78	1.73	1.84
departure	0.07	0.12	0.07	0.00	0.00	−0.05	0.06
LAR log luminous flux	1.39	1.89	1.94	1.76	1.83	1.87	1.87
departure	−0.37	0.13	0.18	0.00	0.07	0.11	0.11
Mean departure (from I.C.I. photopic curve, relative to the value of 542 mμ).	−0.12	0.15	0.18	0.00	0.02	−0.03	0.07

Filter *F* at 542 mμ was used as standard. It is seen that the present data are in agreement with the curves in showing a maximum sensitivity in the region of 560 mμ at a high level of brightness, as compared with the maximum near 510 mμ for the low brightness data of Fig. 12.

DISCUSSION

Data on dark adaptation, light adaptation, and certain other phenomena support the conclusion[11, 12, 13] that the principal *B*-wave of the human retinal response is scotopic in origin. The present experiments provide two more pieces of evidence for this conclusion. First, the electrical data on spectral sensitivity of the dark-adapted eye seem to parallel the scotopic matching data for all except the blue region of the spectrum. Second, the electrical responses of the light-adapted eye fail to reveal any shift from a scotopic to a photopic spectral sensitivity curve. During the electrical experiments, the observers have frequently commented on the extreme brightness of the red lights as compared with the blue ones, even though the electrical responses are of similar magnitudes. This is interpreted to mean that the subjective effect of the flashes is here determined by the photopic system, while the principal electrical effects are scotopic in character no matter how bright the stimuli or how high the level of light adaptation.

It should perhaps be emphasized again that approximately the same filter combinations were used in all four parts of these experiments. Hence it may be concluded that the relationships among the four spectral sensitivity functions are not influenced to any marked degree by differences in band width or other artifacts which might otherwise have played a part. This point is of particular significance because of the fact that the psychophysical data on scotopic sensitivity were obtained at a level about 5.6 log units lower than that of the remaining stimuli. Furthermore, the psychophysical data for scotopic and photopic sensitivity are in good agreement with those of earlier investigators. This fact is taken to mean that no serious error has been made in the basic computations of scotopic and photopic values of luminous flux.

The relatively great effectiveness of blue light for arousing a retinal response is an unexpected result of these experiments, if one accepts the rather convincing evidence that the *B*-wave is governed principally by the scotopic visual mechanisms. There is, of course, no *a priori* reason to expect a perfect correspondence between response magnitudes at the retinal level and those measured by psychophysical procedures. We conclude with the generalization that the spectral sensitivity curves measured electrically by the present technique are obviously different from the usual photopic curves; that they resemble the usual scotopic curves more closely; and finally, that they exhibit a somewhat greater sensitivity to stimulation by lights from 400 to 500 mμ than do the usual scotopic curves obtained by psychophysical procedures.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

Multi-Slit Spectrometry*

MARCEL J. E. GOLAY
Signal Corps Engineering Laboratories, Bradley Beach, New Jersey
(Received March 17, 1949)

The factors which limit the recording speed and the resolution of infra-red spectra are reviewed, and methods offering increased overall photometric efficiency are examined.

The attainment of greater efficiency by means of multi-slit arrangements is discussed and the experimental results are described.

INTRODUCTION

THE relative slowness of the far infra-red spectrometric procedures, and the oft associated sacrifice of resolution, are well known, and it has been customary to ascribe these disadvantages to the relative inefficiency of the radiation sources, as well as to the relative insensitivity of the detectors.

There is undoubtedly considerable justification for this view, albeit good excuses can be offered for both the sources and the detectors designed to work in, and in the vicinity of, the relatively "noisy" spectral region of blackbody radiation at room temperature.

On one side the "narrowband" procedures of radio generation and detection, the possibility of which is inherent in coherent amplification, do not appear to be extendable below the millimeter region, unless and until new and revolutionary principles are invented.

On the other side, the atomic phenomena which permit the generation of nearly monochromatic radiation in, and in the vicinity of, the visible region, do not appear to be extendable to the far infra-red region, and even if they were, monochromatic far infra-red radiation would be of little assistance to the far infra-red spectrometrist who is essentially interested in the absorption spectrum of passive substances.

As to detection, the "quantum phenomena" underlying the operation of the photo-electric and photo-conductive cells become indistinct at around five microns, and even if they did exist beyond this wave-length, their very character of untunability would render them ineffective. Thus, the far infra-red spectrometrist is fairly dedicated to the use of broad banded detectors, which may as well be heat engines, for, in the region of blackbody radiation at room temperature, it makes very little difference, *a priori*, what the underlying principle of these detectors is. Their limitations lie not in them, but in the radiation fluctuations between any surface and its background, which are given by the well known expression first derived by Einstein. Any attempt at escaping these limitations implies the requirement of a large cryostat in which not only the detector, but also the complete monochromator, should be housed.

As one pushes towards longer and longer wavelengths, and reaches the so-called micron region just below the radio limit, a rough narrowbanding of the detector becomes possible by utilizing fine wire bolometers in a high vacuum, but, in the salt optics region, there is no escape from the limitations imposed on sources and detectors when monochromators are operated at room temperature.

It will be noted that the mention of the monochromator is unavoidable when qualifying certain statements about the detectors, and while sources and detectors have been blamed for the disadvantages mentioned earlier, the conventional monochromator should be placed squarely between the other two, just as it is in use, when making the line-up of the elements responsible for these disadvantages.

The most objectionable feature of a monochromator in any spectrometric work in which radiation is at a premium is implied by the very name of the device: only one element of the spectrum is utilized at one time, while the bulk of the radiation is wasted on the jaws of the exit slit. When the source of radiation is small, a far more efficient use of the radiation can be made by employing the method now being developed by Fellgett† for astronomical observations, and which consists in determining the auto-correlation function $\langle f(t)f(t+\tau)\rangle_{Av}$ of the radiation amplitude received by means of an interferometer, and in obtaining the spectrum as the Fourier transform in cosines of this function. This permits to obtain information about the complete spectrum in the time needed with the spectrometer to obtain one section of the spectrum corresponding to the slit width, when the entire source image is accepted at the monochromator entrance.

This procedure bears a resemblance to Michelson's use of the visibility curve, although the circumstance that Michelson's visibility was essen-

* The work reported in what follows was carried out in collaboration with, and at the plant of The Eppley Laboratory, Inc.

† Private communication to the writer.

tially the auto-correlation function referred to some carrier in the center of the narrow spectrum observed, conferred a vectorial character to the visibility; this served to restrict the use of the visibility, obtained visually, and treated as a scalar quantity, to spectra which were symmetrical, or nearly symmetrical, with respect to the carrier. It must also be recalled that Michelson worked with visible light and placed emphasis on the accuracy yielded by his method. The advantage in S/N ration yielded by the interferometric method in the far infra-red, and which does not exist, basically, in the visible region, has been first suggested by Fellgett, to the writer's knowledge.

Fellgett's method is essentially applicable to small sources, and the area of the collimating optics must be divided by the number of elements in the spectrum in order to yield the equivalent permissible source area, from which the radiation can be utilized. If the wave-length of the spectrum studied is equal to a number of times the wave-length spread of the spectrum, the use of zone plate with that same number of zones can permit the utilization of a correspondingly larger source, but even with this modification, the effective radiant area allowed is in general smaller than the available field of the usual infra-red spectrometric optics, which, as will be shown shortly, constitutes the practical limit to the utilizable illuminated area at the entrance of an infra-red spectrometer. As the laboratory infra-red spectrometist utilizes very simple man made sources, he would like, naturally, to be limited by his optical system and to utilize its field fully. Towards this end different arrangements can be had, in which the dispersing monochromator is used as the main stem on which certain new elements are grafted, and as a result of which the spectrum is obtained immediately by the conventional spectral scanning method.

If a balance sheet is made of the radiation values involved in a conventional refracting spectrometer, we begin with a source of blackbody radiation in which some hundreds of watts of electrical power are dissipated. Of the power radiated by the source a matter of watts are focused on the jaws of the single entrance slit of the monochromator while a matter of milliwatts are permitted entrance, out of which a fraction of a microwatt of the desired radiation reaches the detector, through the exit slit. If the wave-length of the desired radiation is shorter than 10 microns, the width of the slits may be made of the order of the wave-length times the numerical aperture without increasing the recording time beyond the presently accepted measure, but if diffracting optics are used, or if the spectral region beyond 10 microns is studied, a heavy premium is placed on the available radiation intensity, and a sacrifice of resolving power must be made if inordinately long recording times are to be avoided. The necessary increased width of both entrance and exit slits is responsible for this loss of resolving power. However, if certain precautions are taken, which form the subjective matter of this report, an increased radiative output can be secured, without loss of resolving power.

In order to visualize readily the general procedure involved, consider that very enlarged entrance and exit slits, which are the image of one another at what will be termed the desired radiation, are subdivided into n slits, of roughly equal width, slits, a_1, a_2, etc. a_n at the entrance, which are imaged onto slits b_1, b_2, etc. b_n at the exit. The desired radiation will therefore have n paths through the "monochromator" (the term will be retained, with an enlarged sense) and, were the passage of the radiation confined to these n paths, an n-fold increase in radiative output would be obtained, consistent with the resolving power which would be realizable if only slits a_1 and b_1 were employed. However, the unwanted radiation entering slit a_1 and passing through exit slits, b_2, etc., b_n would also be measured, and there would be $n(n-1)$ slit combinations defining as many paths for the unwanted radiation. This would result in a loss of the resolving power unless the effect of this radiation were somehow neutralized by the detector and its associated circuit.

One general method of effecting this neutralization, borrowed from communication engineering, could consist in modulating sinusoidally the aperture of the n entrance slits at equally spaced frequencies, $f_1+m\Delta f$, and to modulate sinusoidally the aperture of the n exit slits at the frequencies $f_2+m\Delta f$. In these circumstances, the desired radiation passing through pairs of slits characterized by the same m would have a component of frequency f_2-f_1, and if it is arranged that this frequency does not coincide with any of the individual modulating frequencies, or with any of the sum frequencies, all the radiative components characterized by the same frequency, f_2-f_1, could be detected and yield a measure of the desired radiation, and of no other.

Actually, more information about the spectrum would be contained in the output signal of the detector, for if the slits are properly spaced one from the other, and account is taken of departure from the condition of minimum deviation, the radiation of a different wave-length entering through slit a_m and passing to the detector through slit b_{m+s} would contain a component characterized by the frequency $f_2-f_1+s\Delta f$, which could also be detected and form a measure of the radiation of this different wave-length.

The more general scheme just sketched would require a more elaborate arrangement for processing and presenting the information contained in the

detector signal. If, however, the recording of one spectrum element at a time is merely aimed at, as was the case for the work reported here, the more general scheme of sinusoidal modulation (or "on"-"off" modulation at a constant repetition rate) mentioned above can be replaced by another, more efficient, scheme which will be described below.

ROTATING TWO-DIMENSIONALLY PATTERNED GRATINGS. THEORETICAL CONSIDERATIONS

If a multiplicity of juxtaposed slits are to be modulated, it will be obviously impractical to effect this modulation by means of movable jaws or vanes, and the use of rotating disks suggests itself almost immediately. Furthermore, if the circumferentially variable apertures of a multiplicity of outer concentric annular portions of these disks are kept within the allotted slit width, these annular portions can be crowded next to each other, and the fixed slits entirely dispensed with and replaced by a mask with an aperture corresponding to the utilizable portion of the monochromator field. Thus, virtual slits are formed by the unmasked annular portions. In order to be self-supporting, while having no unused space between the successive modulating rings, these disks should be either re-enforced with radial ribs, or supported on some material which is transparent to the infra-red radiation, or might even be formed by grooving a transparent material so as to cause the radiation passing through the grooved sections to be refracted out of the optical path.

Sinusoidal modulation was not thought practical for three reasons. First, it would be difficult to inscribe it either in unsupported disks, or in thin metallic films backed by a salt disk, and made as indicated above. Second, a large amount of gearing, or a correspondingly involved computer, would be needed to inscribe the hundred or so modulating rings which could be accommodated in the field of most any monochromator. And third, a large amount of useful signal amplitude would be lost, one half in the sum frequency components if the difference frequencies were utilized, the remaining half being further whittled down to $1/8\sqrt{2}$ because of the necessity of fitting a sine curve within the allotted annular space on the disks. This can be readily realized by considering that if the entrance and exit ring (slit) width are unity, the modulation amplitudes will be $\frac{1}{2}(1+\cos 2\pi f_a t)$ and $\frac{1}{2}(1+\cos 2\pi f_b t)$ respectively, which yield the useful output amplitude $\frac{1}{8}\cos 2\pi(f_a-f_b)t$. Thus, in terms of the eventual output signal, the remaining power would be 1/128 of the original available power. Substituting for each sinusoidal modulation the more efficient on-off modulation at a constant repetition rate, the qutput power' would be increased from the figure just given to either $2/\pi^4$ or to 1/48, depending upon whether the first harmonic only, or all harmonics are utilized by the detector circuit, but this procedure would undoubtedly lead to difficulties owing to the presence of unwanted cross-modulation harmonic products between pairs of non-corresponding slits.

TABLE I.

	Entrance slits		Exit slits	
	1st half-period	2nd half-period	1st half-period	2nd half-period
1st slit	0 1 0 1	0 1 0 1	1 0 1 0	0 1 0 1
2nd slit	0 0 1 1	0 0 1 1	1 1 0 0	0 0 1 1
3rd slit	0 1 1 0	0 1 1 0	1 0 0 1	0 1 1 0
4th slit	0 1 0 1	1 0 1 0	1 0 1 0	1 0 1 0
5th slit	0 0 1 1	1 1 0 0	1 1 0 0	1 1 0 0
6th slit	0 1 1 0	1 0 0 1	1 0 0 1	1 0 0 1

With the method of binary modulation which was utilized instead, the desired radiation is subjected to a strong on-off modulation, of which the first harmonic alone is utilized, while the upper odd harmonics are filtered out for reasons which will appear shortly, and in that system the energetic contribution from any pair of corresponding slits to the output signal is equal to $1/2\pi^2$ of the available radiative energy (which is proportional to the square of the radiative intensity or heat flow) through these slits. This fraction compares favorably to the fraction $2/\pi^2$ which is obtained in the case of single slits with the chopping beam method, when the first harmonic alone is utilized, for, on an amplitude basis, it means that *each* corresponding pair of slits yields one half of the radiative output otherwise yielded by single slits with the chopped beam method.

In the binary modulation system, the on-off modulation of every entrance and exit slit is based on the division of a complete period into 2^n parts. A simple instance of this modulation is given in Table I, where 0's and 1's are utilized to indicate whether the radiation is blocked or allowed to pass through the slits.

The examination of the scheme just illustrated reveals that the desired radiation passing through two corresponding slits, is always blocked either at the entrance or at the exit, during the first half-period, while it is allowed to pass through the monochromator for half of the time during the second half-period. On the other side, unwanted radiation passing in and out the 1st and 4th slits, or *vice versa*, and likewise for the 2nd and 5th slits, and the 3rd and 6th slits, is blocked at all times either at the entrance, or at the exit. Lastly, the radiation passing through any other non-corresponding pairs of slits, such as, for instance, the 4th entrance slit and the 2nd exit slit, will be allowed to pass for one quarter of the time during each half-period.

It appears, therefore, that with this scheme the overall radiative output of the monochromator will contain a component which is the desired radiation and which is strongly modulated, being off during the first half-period while being passed through three corresponding slits, on the average, during the second half-period, while all the unwanted radiation will be present the same fraction of the time, on the average, during both half-periods. If, therefore, a radiation detector is employed which yields a square wave output in response to a suddenly applied radiative input, a measure of the desired radiation will be obtained by using circuits which measure the difference of the integrated signals during both half periods.

If we designate the modulation of the mth entrance and exit slits by the functions of time $f_m(t)$ and $f_m^*(t)$, respectively, these functions assuming the values unity for passage and zero for obstruction of the radiation, and no other values, and normalize the time to the value unity for a complete period, we may express the properties of the particular functions shown in the expression above by the relation:

$$\int_0^1 f_m(t) f_n^*(t) p(t) dt = \delta_{m,n},$$

$$\delta_{m,n} = \tfrac{1}{4} \quad \text{for} \quad m = n,$$

$$\delta_{m,n} = 0 \quad \text{for} \quad m \neq n, \qquad (1)$$

where $p(t)$ designates a weight function representing the action of the detector and its associated circuit and having the values -1 for $0 < t < \frac{1}{2}$ and $+1$ for $\frac{1}{2} < t < 1$.

The conjugate "binary functions" $f_m(t)$ and $f_m^*(t)$, exhibit therefore the orthogonal properties with respect to the weight function $p(t)$ which are expressed by relation (1), and which are utilized to realize a cumulative effect of the various radiation beams of the desired wave-length in the finally processed output signal, while the effect of the unwanted radiation is cancelled out. It will be noted that these orthogonal properties are realized individually for each pair of corresponding or non-corresponding entrance and exit slits, and that, therefore, the illumination of the entrance slits need not be uniform. It will also be noted that these binary functions do not share with the sinusoidal functions the property of yielding information on elements of the spectrum others than the one designated as the desired radiation.

The generation of the functions $f_m(t)$ and their conjugates for any value of n can be accomplished as follows. Consider the pattern $\begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix}$ and perform with this pattern the iteration process, symbolized by the pattern itself, which consists in placing another identical pattern at the right of it and below it, and completing the square by placing in the lower right corner the "re-entered" pattern, i.e., the pattern formed by replacing the 0's by 1's, and *vice versa*, in the pattern above. If this process is repeated three times we obtain, successively:

```
00 0000 0000 0000 0000 0000 0000 0000
01 0101 0101 0101 0101 0101 0101 0101
   0011 0011 0011 0011 0011 0011 0011
   0110 0110 0110 0110 0110 0110 0110
        0000 1111 0000 1111 0000 1111
        0101 1010 0101 1010 0101 1010
        0011 1100 0011 1100 0011 1100
        0110 1001 0110 1001 0110 1001
                  0000 0000 1111 1111
                  0101 0101 1010 1010
                  0011 0011 1100 1100
                  0110 0110 1001 1001
                  0000 1111 1111 0000
                  0101 1010 1010 0101
                  0011 1100 1100 0011
                  0110 1001 1001 0110
```

If the first and fourth row of the third pattern are suppressed, the remaining 6 rows will be recognized as the functions $f_m(t)$ given as an example above, whereas the functions $f_m^*(t)$ can be derived by re-entering the first half of each function. It will also be readily verified that if the first and ninth row of the fourth pattern are suppressed, the remaining rows represent so many binary functions which, together with their conjugate functions, satisfy relations (1).

Taken individually, the binary functions just illustrated will be those actually used, although their order within the two dimensional pattern will be changed for reasons which will appear shortly. However, as an introduction to the more complicated computer which will be illustrated later, Fig. 1 illustrates a simple computer which will generate the pattern of 16.

This computer consists of an 8-stage binary counter, C_1 to C_8, 4 gates, G_1 to G_4, each of which delivers a "one" to R whenever both counter stages with which it is associated are in state "one," and a re-entry circuit R, the output of which is a "zero" or a "one" depending upon whether the number of "ones" registered by the gates is even or odd.

Proceeding from left to right, and setting the counter at zero initially the output of the re-entry circuit is written for the first symbol of the first line, 1 is added to the counter, the next symbol is written again in accordance with the re-entry circuit output, and so on, until the 16 symbols of the first line have been thus generated and written down. Passing to the first symbol of the second line, 1 is added again, thus putting the counter in state 0 0 0 1 0 0 0 0 and the symbol of the second, third, etc., lines are generated as for the first. It is readily seen that, except for stages C_4 and C_8, the two upper quarters and the lower left quarter of

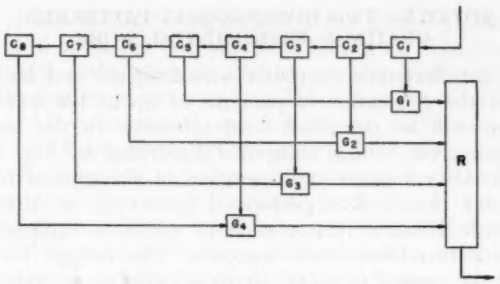

FIG. 1. Block diagram of basic computer.

the pattern of 16 will be generated with identical sequences of states of the various counter and gates, and inasmuch as the output of gate G_4 is 0 for all three quarters, these will be, in fact, identical. Conversely both counters C_4 and C_8 will be in state 1 for the lower right quarter, so that this last quarter will be the re-entry of the pattern of the first three quarters.

It is therefore apparent that the action of the computer corresponds exactly to the passage from a pattern of 8 to a pattern of 16 by the iterative process indicated earlier, and symbolized by the simple pattern $\begin{matrix}0\,0\\0\,1.\end{matrix}$

It follows, by a simple recursion argument, that the pattern formed by the computer illustrated in Fig. 1, is identical with the pattern of 16 shown above, and that any similarly constructed computer with a $2n$-stage binary counter, n gates and a general re-entry circuit will generate the pattern produced by iterating the basic $\begin{matrix}0\,0\\0\,1\end{matrix}$ pattern $n-1$ times with itself.

It has been seen earlier that the postulation of a weight function $p(t)$ which has the values -1 and $+1$ for the first and second-half of a period, respectively, corresponds to the assumption of a far infra-red detector which yields a square wave electrical output in response to a square wave radiative input. As all far infra-red detectors exhibit more or less pronounced departures from this ideal response, their actual response must be analyzed into a fundamental component and odd harmonics of various amplitudes and variously shifted in phase. On the other side, the various unwanted radiative outputs of a monochromator outfitted with disks ruled in accordance with patterns of the type shown above are represented by the various products $f_m(t)f_{m+s}{}^*(t)$, which are neither constant nor identical in both halves of the period. These two circumstances would, in general, prevent the quantities $\delta_{m,n}$ from vanishing when $m \neq n$, which is the condition found earlier for obtaining a measure of the desired radiation which is not affected by the radiation in the rest of the spectrum.

In order to achieve reasonable assurance that only the desired radiation is measured for any one setting of the monochromator, the following three steps were taken.

First, the fundamental harmonic alone of the detector signal was utilized, the higher odd harmonics being filtered out. This served to eliminate any concern about the effect of the amplitude and phase of these harmonics, so that $p(t)$ could be replaced simply by a sine function, while the lower sensitivity achieved in this manner amounted to the fraction $2\sqrt{2}/\pi$ of the sensitivity theoretically available from a detector with a square wave response.

Second, the entrance apertures of the monochromator were illuminated as uniformly as possible.

Third, the binary functions illustrated above were transposed line by line, and certain ones of them were re-entered, in accordance with the scheme which will be described briefly below, and more completely in a separate article.*

The two first steps mentioned above permit to re-write as follows the condition that the measurement of the desired radiation be unaffected by other radiative components:

$$\int_0^1 \sum^m f_m(t)f_{m+s}{}^*(t)\sin 2\pi t = 0. \qquad (2)$$

In the expression above $\sin 2\pi t$ is the weight function which characterizes a detector-amplifier-rectifier combination which yields a measure of the fundamental harmonic only of the detected signals, and this weight function is applied to the *sum* of all the unwanted radiative components which are characterized by the same wave-length, and each of which emerges from the monochromator through the modulated slit which is s slits distant from the exit slit corresponding to the entrance slit for this particular component.

Of the two functions under the integral sign of condition (2), one, $\sin 2\pi t$, is a continuous function of time, while the other is a discrete, discontinuous function restricted to the values 0 or 1 in anyone of the 16 (or 2^n) intervals into which a period is divided. As the only property shared by continuous and discrete discontinuous functions is symmetry, condition (2) will be satisfied if the sum functions, $\sum^m f_m(t)f_{m+s}{}^*(t)$ satisfy one of the two following conditions of symmetry, which are satisfied by $\sin 2\pi t$, except for the change of sign:

A. The sum functions have the same value when $t=\frac{1}{2}-a$ and when $t=\frac{1}{2}+a$, i.e., the sum functions are symmetrical with respect to the epoch $t=\frac{1}{2}$.

* This other article will appear in a forthcoming issue of the J. Math. and Phys. (M.I.T.).

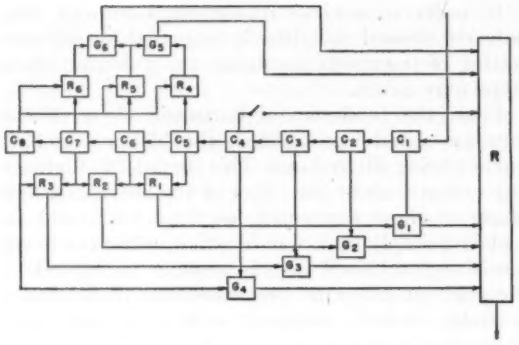

FIG. 2. Block diagram of typical computer.

B. The sum functions have the same value when $t=a$ and when $t=\frac{1}{2}+a$, i.e., the sum functions are identical over both half-periods.

(To the functions which satisfy one or the other of these two conditions may be added the functions which are the sums of two functions satisfying conditions A and B, respectively, but these will not play a role in what follows.)

It can be verified that the binary functions shown above do not satisfy conditions A nor B, and the generation of suitable functions permitting the realization of conditions A or B was accomplished by replacing the simple type of computer illustrated in Fig. 1 by the more elaborate type of computer in Fig. 2.

As the underlying theory of this type of computer will be discussed in another article, Table II is merely designed to illustrate its arithmetical operation. The output of the re-entry circuits R_1, R_2 and R_3, which are represented by the last three figures of the progressive cycle number of the rows, play the part formerly played by the output of stages C_5, C_6 and C_7, and the re-entry of gates G_5 and G_6 serves to determine the re-entry of the 4th, 7th, 10th, and 13th rows.

The single switch shown connected to stage C_4, in the block diagram of the computer, is closed when it is desired to generate the conjugate f^* functions. This is accomplished by adding an additional re-entry to R whenever C_4 is in state 0, which causes the re-entry of the first half of the f functions.

The sum functions are calculated, for the first half of the pattern, by counting how many 1's are s lines above a 0 (corresponding to a 1 of the f^* functions) in each column, and, for the second half of the pattern, by counting how many pairs of 1's are s elements apart. The desired A or B symmetry can be observed in the four lines giving the values of the sum functions for the values $s=1$, 2, 3, and 4, which have been written below the pattern.

ROTATING TWO-DIMENSIONALY PATTERNED GRATINGS. EXPERIMENTAL WORK

An electronic computer was designed and built for the generation of patterns of up to 128×128, but will be described with reference to the less bulky, yet typical computer illustrated by Fig. 2. In this computer the operation of the general re-entry circuit R is performed by means of Automatic Electric relays, each of which is equipped with two break-make contacts. The output lead of the general re-entry circuit is either grounded or left open, and the energization or release of any one relay serves to pass from one connection to the other.

One 6SN7 double triode is utilized for each counter stage, and each gate. Each counter stage consists of a conventional double stability circuit with memorizing capacitances between each plate and the opposite grid. Buffers not indicated in Fig. 2 are provided to insure the reliable operation of the counter.

The re-entry circuits R_1 to R_6 are actually phantom circuits, indicated in the block diagram in order to simplify arithmetical discussion, but having no physical counter parts. For instance, gate G_1 must be conductive whenever counter stage C_1 is in state 1 while stages C_5 and C_6 are in state 0 and 1, irrespectively. This was realized by connecting one of the grids of the double triode, through three 7 megohms resistors, to the plate of C_1 which is non-conductive when C_1 is in state 1, and to the plates of C_5 and C_6 which are non-conductive when these stages are in states 0 and 1, respectively. The other

TABLE II.

				Column ordinal number (reading downwards)			
				0000	0000	1111	1111
				0000	1111	0000	1111
				0011	0011	0011	0011
Row ordinal number	Row progressive cycle number	G_5	G_6	0101	0101	0101	0101
0000	0000	0	0	0000	0000	0000	0000
0001	0001	0	0	0101	0101	0101	0101
0010	0011	0	0	0110	0110	0110	0110
0011	0010	1	0	1100	1100	1100	1100
0100	0110	0	0	0011	1100	0011	1100
0101	0111	0	0	0110	1001	0110	1001
0110	0101	0	1	1010	0101	1010	0101
0111	0100	1	1	0000	1111	0000	1111
1000	1100	1	1	0000	1111	1111	0000
1001	1101	0	1	1010	0101	0101	1010
1010	1111	0	0	0110	1001	1001	0110
1011	1110	0	0	0011	1100	1100	0011
1100	1010	1	0	1100	1100	0011	0011
1101	1011	0	0	0110	0110	1001	1001
1110	1001	0	0	0101	0101	1010	1010
1111	1000	0	0	0000	0000	1111	1111
Sum functions		$s=1$		4444	3333	3333	4444
		$s=2$		4343	4343	4343	4343
		$s=3$		1441	4334	4334	1441
		$s=4$		3333	3333	3333	3333

grid was similarly connected to the same C_1 plate and to the other two plates of C_5 and C_6. The bias of both grids was provided by connection through a 3 megohm resistor to a -275-volt supply, while a $+145$-volt supply was utilized for the counter plates. Both plates of the gating tube were connected together and to a $+90$-volt supply through the winding of the corresponding relay in R. Thus two triple gates are formed, one or the other of which is conductive when, and only when, counter stages C_1, C_5 and C_6 are in states 1, 0 and 1 or 1, 1 and 0, respectively.

Likewise, gate G_5, for instance, is formed of two triple gates with a common output, one of which is conductive when, and only when, counter stages C_5, C_6 and C_8 are in states 1, 1 and 0, respectively, while the other is conductive when, and only when, these same counter stages are in states 0, 0 and 1.

The ruling engine employed for ruling the gratings was built around an accurately machined main worm gear on which the disks ruled into gratings are fastened. The prime moving shaft is motor driven and geared to the main worm gear through a 512:1 reduction ratio.

At each turn of the prime moving shaft a photocell is flashed once and delivers an electrical impulse to a scaling circuit placed ahead of the computer proper.

The ruling tool is allowed to make a mark on the disks, or is lifted from the disks, by an electromagnet, in accordance with the instructions received from the computer, the general re-entry circuit of the computer either completing to ground the energizing circuit of the electromagnet, or opening it.

While the main worm gear and the disk being ruled rotate slowly, the ruling tool is traversed extremely slowly from the edge towards the center of the disk, the distance traveled for one rotation of the disk corresponding to the width of one slit, and being roughly equal to the width of the ruling tool. Thus the disks are ruled along spirals rather than along concentric circles, but this introduces no appreciable difference in their operation. The traversing motion of the ruling tool is accomplished by means of an accurately cut worm gear engaged in a matrix to which the ruling tool and the lifting electromagnet are attached. This worm gear is geared to the prime moving shaft through two worm and worm gear speed reducers.

In the work done so far rock-salt disks 3 in. in diam. and $\frac{3}{16}$ in. thick have been used for the gratings. The original plan had been to groove the polished disks so as to throw out of the spectrometric path the radiation falling on the grooved portions which were to act thus as opaque portions. A more practical scheme, suggested by Mr. A. Dennett, of the Eppley Laboratory, was adopted subsequently, which consisted in evaporating in vacuo a layer of gold on the polished rocksalt disks, and in removing transparent strips by means of a relatively soft tool. Eventually a layer of gold black was added on the metallic gold layer in order to insure that unwanted radiation reflected internally within the spectrometer would not cause unwanted effects.

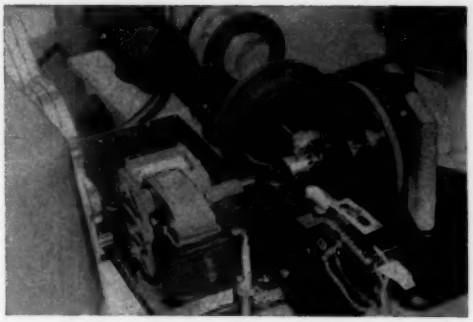

FIG. 3. Rotating two-dimensionally patterned gratings mounted on infra-red spectrometer.

Figure 3 is a close-up photograph illustrating the mounting of two "mated" gratings, ruled in accordance with the f and f^* functions of a 64×64 pattern, in place of the conventional slits in a Perkin Elmer spectrometer. The gratings are rotated at 10 c.p.s. by a 2-pole 60 c.p.s. synchronous motor with a 6:1 reduction gear built in, and are geared with each other by two miter gears. The radiation passes to the entrance portion of the entrance grating on both sides of the thin shaft of the exit grating. Since both gratings have the same diameter, no correction for slit curvature is effected, and, for this reason, the effective slit height was reduced from 13 mm for the single slit to 8 mm for the 64-slit arrangement shown. This was done by means of a mask with an 8×8 mm opening, placed just inside of the entrance grating, a corner of it is visible in the photograph. A small 45° mirror reverses the dispersed spectrum on the exit grating so as to permit the perpendicular arrangement of the gratings. The out going radiation is focused on a pneumatic detector with a $\frac{1}{8}$-in. diam. sensitive area by means of the AgCl lens shown. A photocell and a small lamp bulb are shown in the foregound, both blackened except for the small portions of their glass envelopes facing each other. Their light path is periodically opened and closed by the rotating sector attached to the entrance grating shaft, and the square wave signal thus produced in the photocell serves to trigger a two-stability circuit, the output of which is mixed with the amplified and filtered output of the detector, and rectified, so as to produce a d.c. output which is proportional to

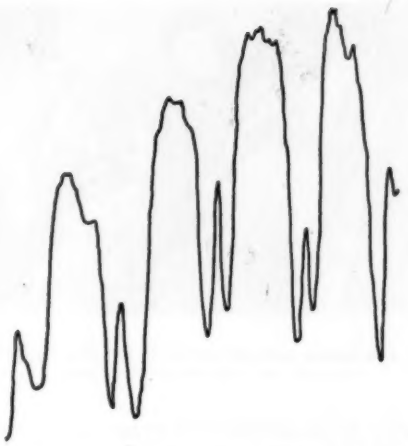

FIG. 4. Tracing of ammonia spectrum record taken with rotating two-dimensionally patterned gratings. Wave number approximately 850 cm^{-1}. Equivalent slit width 0.126 mm.

the on-off fluctuations of the desired radiation component.

Since roughly one-third of the radiation was blocked by the thin shaft of the exit grating, while 8/13 of the normal slit height was utilized and additional losses occurred at the four grating surfaces, a signal increase of only ten was expected over the conventional slit arrangement, as against the increase of 32 theoretically obtainable from 64 slits. This tenfold increase was roughly verified by masking the 1½-in. diam. AgCl lens with a shield having a ½-in. hole, and noting that the signals obtained were equivalent to those obtained with the conventional slits. Figure 4 illustrates a portion of the well-known ammonia spectrum obtained in this manner.

CONCLUSIONS

The work done so far and the results obtained have demonstrated the practical feasibility of multi-slit spectrometry. Further technical elaboration will serve to render unneeded the compromises resorted to in order to expedite the original work. For instance, the correction for slit curvature could be accomplished by utilizing two gratings, of the proper dimensions, with a ⅔-diameter ratio, and by repeating the modulating pattern twice on the smaller grating and three times on the longer one, so as to insure proper meshing of the image of the entrance pattern with the exit pattern. It is believed, therefore, that the radiative output of a spectrometer can be increased by half the number of slits permitted by the optics, i.e., by the ratio of one quarter the utilizable field width divided by the width of the slit it is desired to use. Thus an appreciable signal increase can be yielded by conventional spectrometers, while the application of the Schmidt optical principle to infra-red spectrometry will yield additional increases.

In order to discuss the effect these increases in radiative output will have on S/N ratio, recording speed and resolving power, account must be taken of the detector utilized. In the experimental work reported here a pneumatic detector with a ⅛-in. diam. sensitive area was utilized throughout, and it will be readily seen that, with no change in the detector, an n-fold increase in radiative output will yield a like increase in S/N ratio, and an n^2 fold increase in speed. If, however, the size of the utilizable field of a spectrometer becomes so increased that an increase in detector area becomes necessary, the sensitivity of the pneumatic detector will decrease slightly at first, until certain optimum condition are reached, which are determined by the lowest gas conductivity available (Xenon). Beyond diameters of the order of 1 cm the sensitivity of the pneumatic detector can be expected to decrease in inverse proportion to the square root of its area, and to be limited to a ceiling equal to 57 percent $(1/\sqrt{3})$ of the sensitivity of an ideal detector of the same area. Because of these considerations it is believed that it will be possible to extend the infra-red spectrometric methods to the radio region (1 mm) while realizing the resolving power inherent in the optical system utilized.

ACKNOWLEDGMENT

The writer wishes to express his appreciation to the Eppley Laboratory Staff, and more particularly to Dr. Marion Eppley, Mr. Roy Anderson, Mr. Aleck Dennett, and Mr. Karl Falkenholm, for the splendid cooperation and the enthusiasm with which they have made the work reported here possible.

The writer would like also to express his thanks to his wife, who has made many valuable editorial comments.

The Plastic Deformation and Crystal Orientation of Thallium Halides*

ALEXANDER SMAKULA AND MYRON W. KLEIN
Engineer Research and Development Laboratories, The Engineer Center, Fort Belvoir, Virginia

(Received March 10, 1949)

The plastic deformation of cubic thallium halide crystals produced under pressure of a cone perpendicular to the crystal surface, shows two kinds of very sharply limited, raised patterns, namely, the "superficial" and the "transmitted." In both of these patterns, the shape, symmetry, and position depend on the lattice orientation of the crystal with respect to the observation surface. The first appears around the press point in the form of radial wings. The second develops on the opposite side of the press point and has a square or a rhombic shape. Characteristics of the patterns are explained by the localized distortion of the single crystal around the press point into an aggregate of small crystals. The glide planes (T) of the small crystals lie in the dodecahedron faces (110), and the glide directions (t), in the cubic faces (001). The superficial patterns can be applied to determine the crystal orientation within a few degrees. Inner strains which are not permanent develop symmetrically around the press point. Under pressure of a cone, sodium chloride crystals show splitting along the dodecahedron faces (110), while under the same conditions, silver chloride crystals give only diffuse patterns.

INTRODUCTION

SUBSEQUENT to the investigations of Rubens[1] natural and synthetic crystals have been used in the form of prisms, lenses, and windows for scientific and applied investigations in the ultraviolet and infra-red spectroscopy. Although alkali halides (i.e., NaCl, KCl, and KBr) have been used for a long time, their high water solubility is a serious disadvantage because exposure to moist air destroys optical surfaces. The resistance to moisture of optical elements made from these crystals may be somewhat improved by heat treatment[2] or suitable coatings, but such improvement is not sufficient to permit their use under average humidity conditions. The use of silver chloride,[3] on the other hand, is limited because of its very low hardness which makes grinding and polishing quite difficult. Crystals of thallium halides have some properties better than those of alkali and silver halides. Their water solubility is less than that of the alkali halides, and they are practically insensitive to visible light. Furthermore, their high refractive indices[4] and high transmission over wide ranges in the infra-red[5] make them desirable for some applications. Because of their greater hardness, the mixed crystals of thallium halides have been developed in the laboratory of the senior author (A. S.) at Carl Zeiss Optical Company, Jena, by R. Koops.[6] These crystals, which are presently known under the code names KRS-5 (=TlBr+TlI) and KRS-6 (=TlCl+TlBr), are finding important applications.[7] The infra-red cut-off for KRS-6 is approximately at 30 microns and for KRS-5 at 40 microns.[8] The refractive indices from 0.589 micron to 24 microns for KRS-6 and from 0.589 micron to 32 microns for KRS-5 have been estimated by Hettner and Leisegang,[9] and from 32 microns to 39.5 microns by Plyler.[10]

In order to make the most efficient use of thallium halide crystals it is necessary to know, in addition to the infra-red optical properties, such mechanical properties as hardness, elasticity, and plasticity. Photoelastic properties of thallium halides have been investigated by West and Makas.[11] Unlike the optical properties, the mechanical properties depend on the orientation of the crystal lattice. This

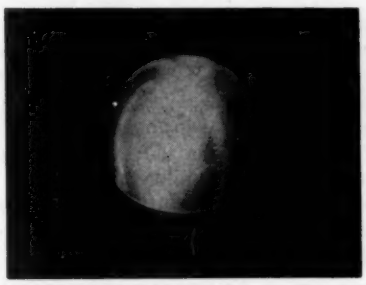

FIG. 1. Crystal hemisphere of KRS-5 for punch patterns before polishing.

* Paper presented at a meeting of the American Optical Society in Detroit, Michigan, October 21–23, 1948.

[1] H. Rubens, Verh. d. D. Phys. Ges. **15**, 108 (1896); **17**, 102 (1911); Wied Ann. **69**, 576 (1899); Phys. Zeits. **4**, 726 (1903); Ann. d. Physik. **26**, 615 (1908); **54**, 476 (1895); **60**, 724 (1847).

[2] H. Rather, Optik **1**, 296 (1946).

[3] Raymond M. Fuoss, Rev. Sci. Inst. **16**, 154 (1945); H. C. Kremers, J. Opt. Soc. Am. **37**, 337 (1947).

[4] H. Schroeder, Zeits. f. Physik. **67**, 24 (1931); T. Barth, Am. Mineral. **14**, 358 (1929).

[5] R. B. Barnes, Zeits. f. Physik. **75**, 730 (1932).

[6] A. Elliot and L. E. Mayes, *German Infrared Devices and Associated Investigations*, CIO Subcommittee Report No. 2, File XXXIII-9, 8–10 (1945); R. Koops, Optik **3**, 298 (1948).

[7] E. K. Plyler, J. Chem. Phys. **15**, 885 (1947); Robert Hofstader, Phys. Rev. **72**, 1120 (1947); W. L. Hyde, J. Chem. Phys. **16**, 744 (1948).

[8] Measurements of M. Czerny see R. Koops, reference 6.

[9] G. Hettner and G. Leisegang, Optik **3**, 305 (1948).

[10] Private communication by E. K. Plyler.

[11] C. D. West and T. Makas, J. Chem. Phys. **16**, 427 (1948); E. Burstein and P. L. Smith, Phys. Rev. **74**, 229 (1948).

FIG. 2. Pressing goniometer with horizontal and vertical circle.

paper contains an investigation of plastic deformation and crystal orientation of thallium halides as an aid in the grinding and polishing of optical elements.

METHOD

The investigation of plastic deformation of crystals is generally made by applying unidirectional tension or compression. This method is tedious for a series investigation because it requires a new sample for each measurement. Therefore, a more convenient so-called "punch-pattern" method which was used in crystallography for the first time by Reusch[12] was employed in this study. The method consists of pressing a needle into the crystal surface. Raised patterns which can be seen under a microscope appear around the press point. The shape and the symmetry of the patterns depend on the crystal orientation.[13] To enlarge the patterns, a cone with an 80-degree apex angle was substituted for the needle. The patterns could be seen, then, by the unaided eye if the applied force was higher than 7 pounds. A special advantage of this method is the possibility of seeing in one picture the deformation of the crystal by forces originating from one central point and covering an angle of 360 degrees.

INVESTIGATION OF THALLIUM BROMOIODIDE (KRS-5) BY USE OF PUNCH-PATTERN METHOD

The thallium halide crystals have a body-centered cubic lattice with the exception of thallium iodide which crystallizes rhombic. The mixed crystals also have a body-centered cubic lattice. According to the method previously described, the punch pattern on a cubic face (hexahedron) (001) should have fourfold symmetry, on a dodecahedron face (110), twofold symmetry, and on an octahedron face (111), threefold symmetry.

In order to study the general shape and symmetry of patterns, and their relation to the crystal orientation, a hemisphere of KRS-5 with a radius of 2.5 cm was prepared with its base parallel to the cubic face. In this way, all possible orientations were exposed for investigation. The hemisphere after having been turned to shape on a lathe is shown in Fig. 1. The rough bands spaced 90 degrees apart around the axis of the hemisphere indicate a difference in behavior of various orientations under turning. The hemisphere was finally polished by conventional optical methods. In order to press the punch into any desired orientation a pressing goniometer was constructed (Fig. 2). The horizontal and vertical setting circles of the goniometer enable one to obtain any desired orientation within an accuracy of 2 degrees. After crystal setting, the load was placed upon the pan, thus pushing the punch into the surface. The conical punch with an apex angle of 80 degrees under a load of about 20 pounds gave the best results. With this load the punch sinks into the surface to a depth of about 0.5 mm and produces a hole of approximately 0.8 mm in diameter. Around the hole appear one to four radial raised wings which might be called "superficial" patterns, the length and symmetry of which depend upon the orientation of the crystal.

SUPERFICIAL PATTERNS

The wings which may be easily seen by the unaided eye were photographed in reflected light. Three series of patterns were produced at various angles, as shown in Fig. 3, with angular steps of 10

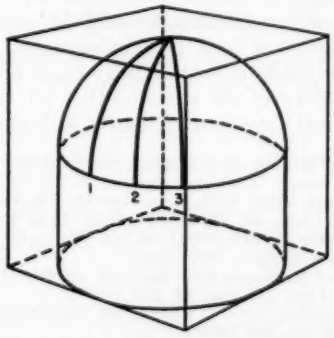

FIG. 3. The model of KRS-5 sphere used for punch patterns. The cube indicates the crystal orientation. The lines marked 1, 2, and 3 indicate the directions in which the three series of punch patterns shown in Figs. 4, 6, and 7 were taken.

[12] E. Reusch, Ann. d. Physik. and Chemie 132, 443 (1867); G. Muegge, Neu. Jahrb. f. Mineral. 1, 72 (1898); A. Johnson, Neu. Jahrb. f. Mineral. 2, 145 (1902); Fortschr. d. Mineral 3, 100 (1913); Jahrb. d. Rad. and El. 11, 248 (1914).

[13] G. Tammann and A. Mueller, Zeits. f. Metallkunde 18, 69 (1926).

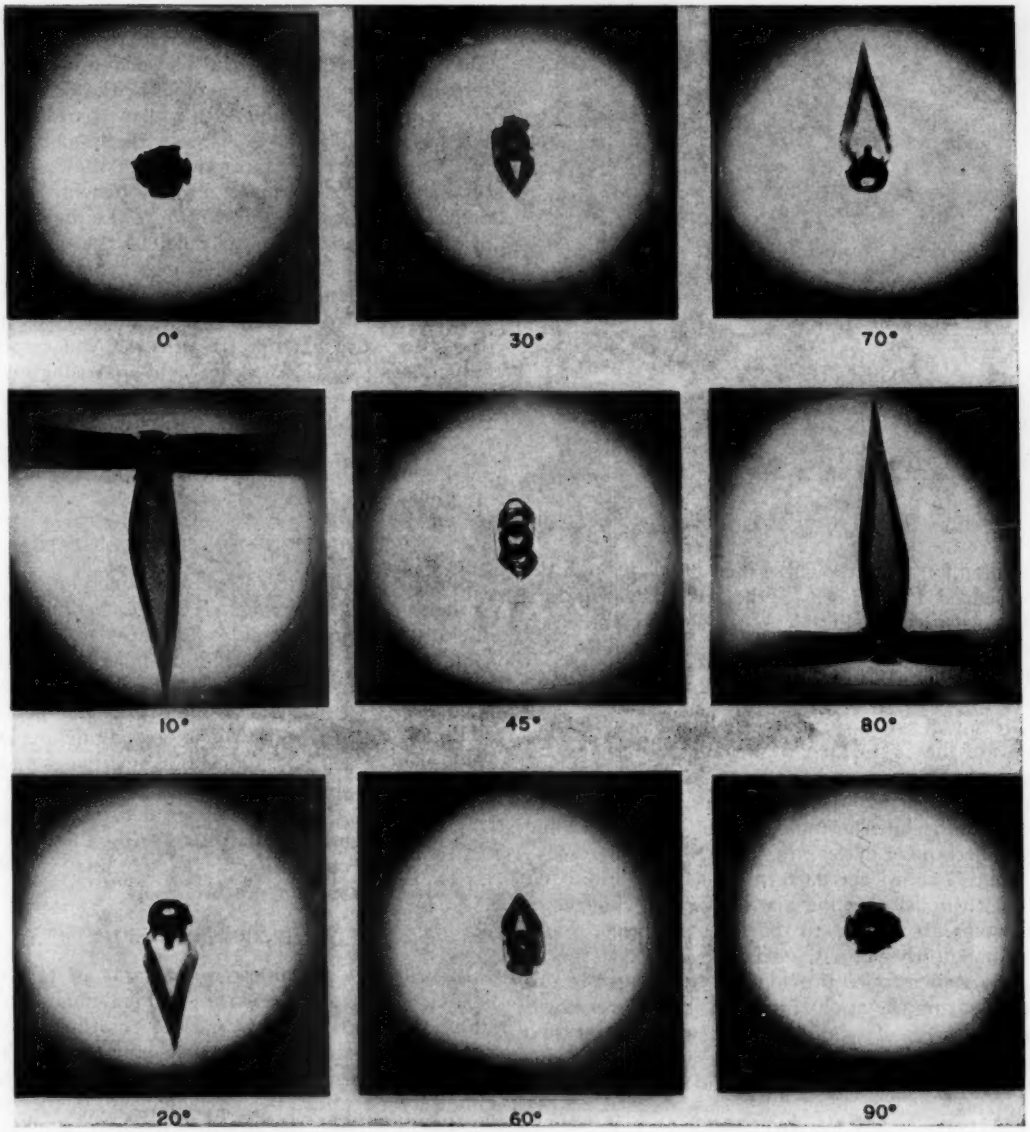

FIG. 4. Punch-pattern sequence from the cubic face (001) to the next cubic face (010). Magnification 5.7.

degrees. The first series was taken from one cubic face (001) to another cubic face (100); the second series from the cubic face (001) through the octahedron face (111) to the dodecahedron face (110); and the third series from the cubic face (001) along the bisector to the base of the hemisphere. Figure 4 shows the first series of patterns. On the cubic face (001) there is a slight indication of fourfold symmetry on the periphery of the press hole. Under the proper lighting conditions one can detect four very long weak wings emanating from the hole perpendicular to the edges of the cube, as shown in Fig. 5. As the angle is increased with respect to the cubic face, three of the original wings disappear. Only the wing which points in the direction of advance remains. With angular change, this wing becomes successively shorter and stronger, and a second wing develops on the opposite side of the hole. At 45 degrees both wings become equal. Continuing along this path the new wing becomes successively longer and the old wing disappears.

Figure 6 shows the sequence of patterns made in

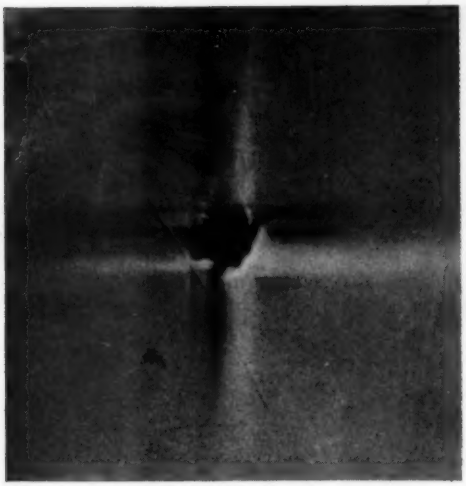

FIG. 5. Punch-pattern wings on the cubic face.

the direction from the cubic face (001) through the octahedron face (111) to the dodecahedron face (110), as shown in Fig. 3. As the angle is increased the two long symmetrical wings become progressively shorter and stronger, while the third wing gradually develops. The angle between the two original wings increases and reaches 120 degrees at the (111) face where all three wings are equal. As further advance is made the third wing becomes longer and disappears entirely on the edge of the cube, while the angle between the two original wings becomes 180 degrees.

Figure 7 shows the pattern sequence as indicated in Fig. 3 under the line marked 2. The changes of the wings are similar to the foregoing Figs. 4 and 6, although no symmetry exists.

As was mentioned previously, wing length is also dependent on the applied load. In Fig. 8 are shown typical curves of wing length *versus* load for two orientations. In both cases, the wing length approaches the maximum as the load is increased.

STRAIN PATTERNS

In addition to the patterns which appear on the crystal surface by pressing, strains are introduced just below the surface. These may be seen between crossed polaroid screens. Figs. 9, 10, and 11 show strain patterns for punch and observation directions (100), (110), and (111). Corresponding to the crystal orientation the strain patterns show 4-, 2-, or 3-fold symmetry but they seem to be less sensitive to crystal orientation.

TRANSMITTED PATTERNS

If crystal samples only a few millimeters thick are used to produce the punch patterns, one obtains on the opposite side of the indentation a pattern in the form of a broad truncated pyramid, called the "transmitted" pattern. The shape of the pattern and its position depend upon the orientation of the crystal. Punching into a plane parallel plate the surfaces of which coincide with the cubic face, produces a raised square pattern in the form of a broad truncated pyramid directly beneath the indentation, as shown in Fig. 12, top. If either or both of the surfaces are not parallel to the cubic face the transmitted pattern becomes rhombic and is displaced laterally from the point beneath the indentation. Pressing on a dodecahedron face produces two transmitted patterns on the opposite cubic faces (Fig. 12, bottom) while pressing on the octahedron face produces three transmitted patterns. Figures 13, 14, and 15 show schematically the relation between the position of the superficial and transmitted patterns, and the crystal orientation.

INVESTIGATION OF OTHER CRYSTALS

Similar punch patterns were obtained on crystals of thallium chlorobromide (KRS-6), thallium chloride, and thallium bromide. Attempts to apply the same method to silver chloride failed. Instead of sharp wings there appeared a diffuse fourfold symmetrical pattern, which seemed to be independent of the crystal orientation (Fig. 16). Punching on a cubic face of sodium chloride in most cases results in fourfold splitting along the diagonals of the cubic face (Fig. 17). This experiment was repeated at a temperature of about 400 centigrade, and the press pattern was similar to that obtained on silver chloride at room temperature.

DISCUSSION

In order to interpret the patterns in terms of behavior of crystal orientation, one must understand the mechanism of the deformations. Previous investigations of plastic deformation under unidirectional tension or compression made on alkali halide crystals, have been explained in the following manner:[14] Under the applied pressure the single crystal is broken into many small parts which glide along special planes and in special directions which are called the glide planes and glide directions.[15] Thus, a plastically deformed region of the crystal is no longer a single crystal but an aggregate of small crystals which are still mechanically connected. In NaCl and other face-centered cubic crystals the glide planes (T) are parallel to the

[14] A. F. Joffé, *The Physics of Crystals* (McGraw-Hill Book Company, Inc., New York, 1928); M. J. Buerger, Am. Mineral. **15**, 174 (1930).

[15] The term "glide" is used to indicate the general process of translation and twinning. See: W. Boas, *Physics of Metals and Alloys* (John Wiley & Sons, Inc., New York, 1947), p. 71.

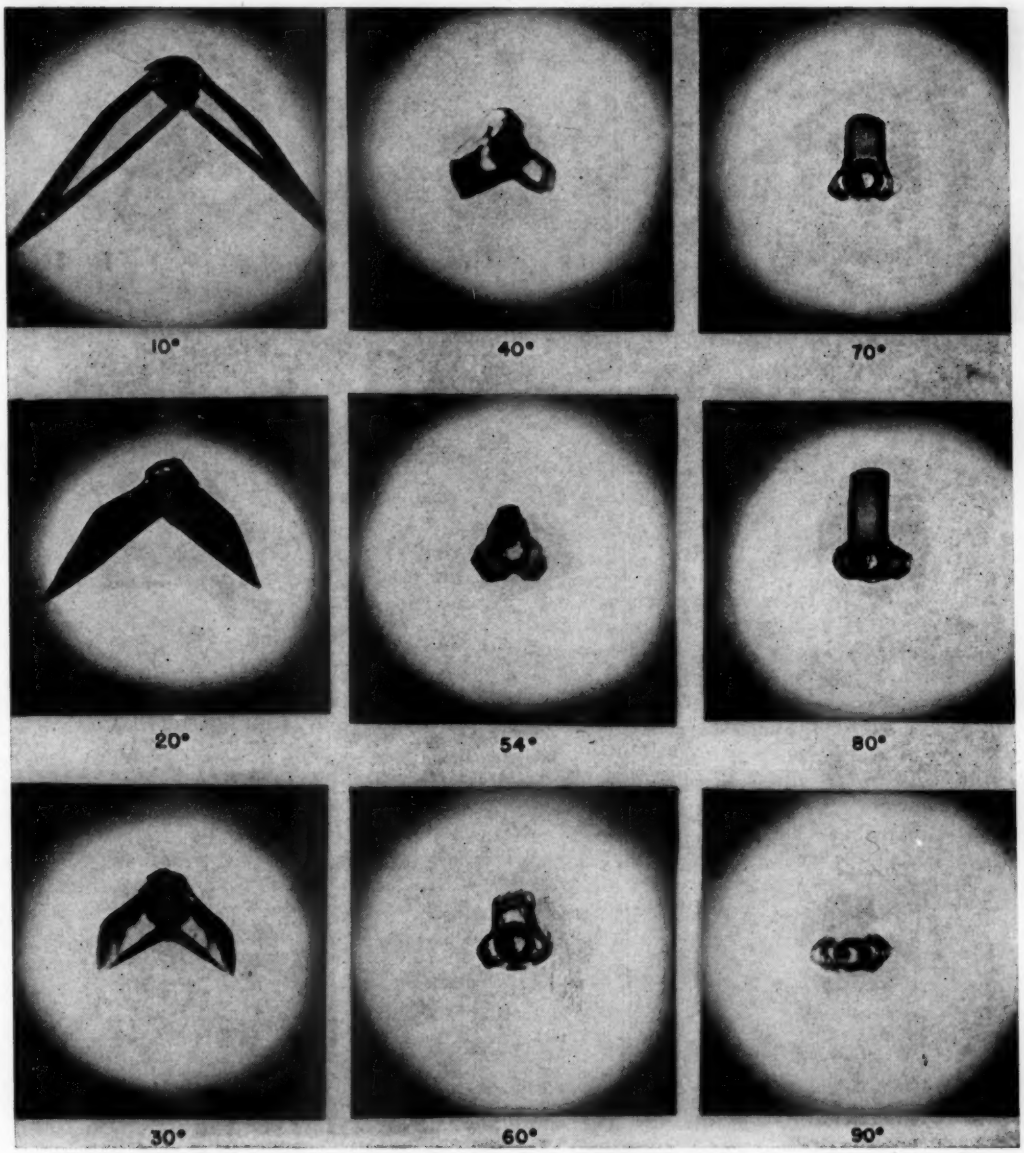

FIG. 6. Punch-pattern sequence from the cubic face (001) through the octahedron face (111) to the dodecahedron face (110). Magnification 5.7.

dodecahedron faces. The glide planes contain positive and negative ions, while the glide directions (t) are in the planes having the densest rows of ions with like charge.

In the case of cubic thallium halides the deformation takes place in translation and twinning. The glide planes and glide directions can be determined from the position and the shape of the transmitted patterns (Figs. 13, 14, and 15). The corners of the transmitted patterns are turned 45 degrees with respect to the cubic face. This position indicates that the material glides only in dodecahedron planes. A line between the press point and the transmitted pattern indicates the glide direction which is toward a cubic face. In thallium halides the glide planes and glide directions are just reversed as they are in sodium chloride. This difference is caused by the various lattice structures.

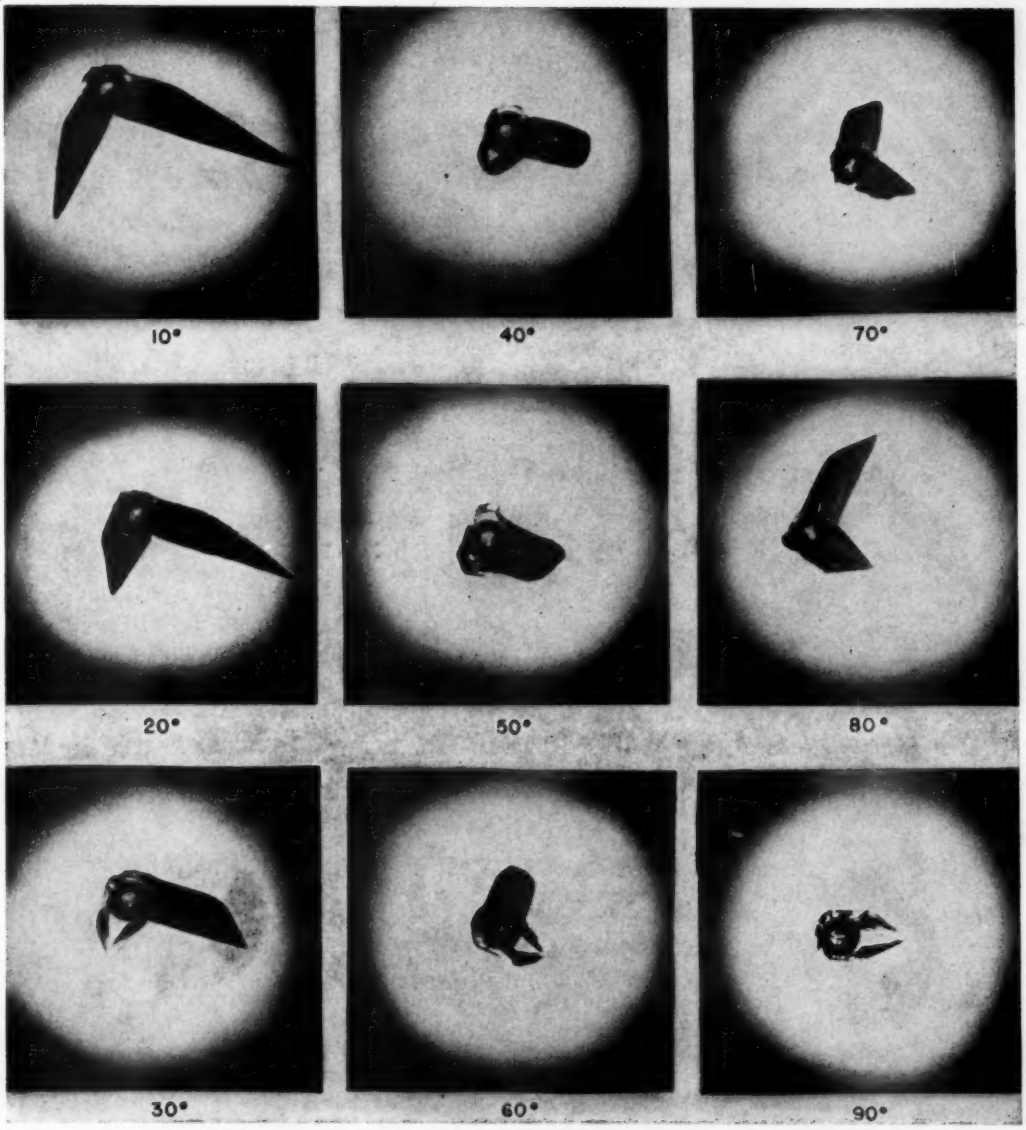

FIG. 7. Punch-pattern sequence from the cubic face (001) along the path midway as shown in Fig. 3, under the line marked 2. Magnification 5.7.

Both crystals are cubic, but while the lattice of sodium chloride is face centered, the lattice of cubic thallium halides is body centered. In both cases, the glide planes contain positive and negative ions and the glide directions lie in the densest row of ions having like charge.

To explain the change of the wing length of superficial patterns the following factors must be considered: (1) the applied force, (2) the glide plane, (3) the glide direction, and (4) the observation plane. If one proceeds from the cubic face to the dodecahedron face (Fig. 4) the length of the wing l is given by the following equations,

$$l = l_0 \cos\varphi \cos\lambda \cos\vartheta,$$

where l_0 = the length of the wing on the cubic face, ϑ = the angle between the glide direction and the observation plane, φ = the angle between the radial force direction and the glide plane (T), λ = the angle between the radial force direction and the

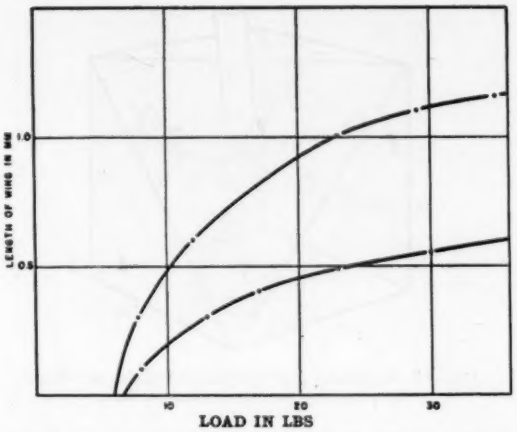

FIG. 8. The influence of the load on the length of the press patterns; lower curve: 110-face; upper curve: surface between 001- and 110-face.

glide direction (l). In this case, $\lambda = \vartheta$. On the cubic face there are $\varphi = 45$ degrees and $\lambda = \vartheta = 0$. By proceeding toward the dodecahedron face φ increased to 90 degrees and $\lambda = \vartheta$ to 45 degrees.

In Fig. 18 are given the lengths of the wing which are obtained by punching a cone with 20-pound force proceeding from the cubic face (001) (0 degrees) to the dodecahedron face (011) (45 degrees), and the corresponding curve computed according to the theoretical equation. In general, the experimental points fit the curve very well. A deviation is found only around the dodecahedron face (45 degrees) where the experimental values are higher than the theoretical. In this case, the wings are irregular and not well-defined. Probably there is in this region an inhomogeneous deformation, where the crystal symmetry is destroyed.

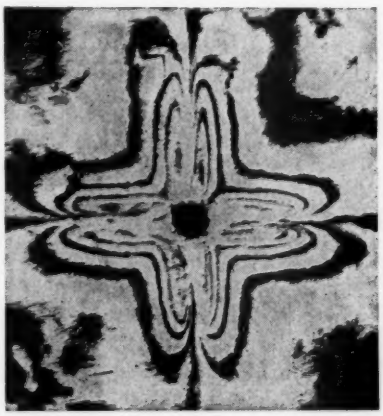

FIG. 9. Inner strains formed by punching on the 001-face of KRS-5.

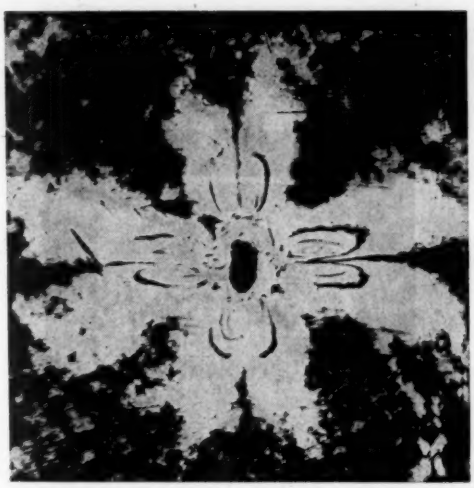

FIG. 10. Inner strains formed by punching on the 110-face of KRS-5.

Knowing the change of the wings with crystal orientation one can use them as a convenient method to determine the crystal orientation. This method is advantageous for thallium halide crystals since the other methods (goniometrical or x-ray) require good surfaces without lattice distortion; however, such surfaces are very difficult to obtain on crystals with high plasticity. The first application of this method to KRS-5 was made by Koops.

The shape of the strain patterns produced by conical punching, observed in the press direction, also depends on the crystal orientation. On the

FIG. 11. Inner strains formed by punching on the 111-face of KRS-5.

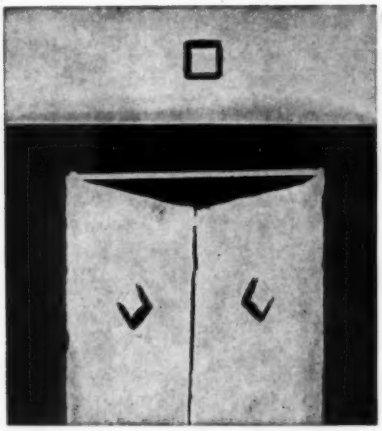

FIG. 12. Transmitted punch patterns on KRS-5. Top: by punching on the cube face; bottom: by punching on the dodecahedron face.

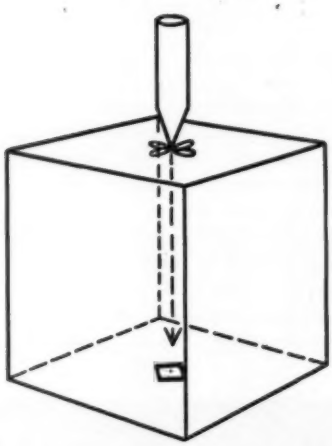

FIG. 13. Superficial and transmitted patterns formed by punching on a cubic face.

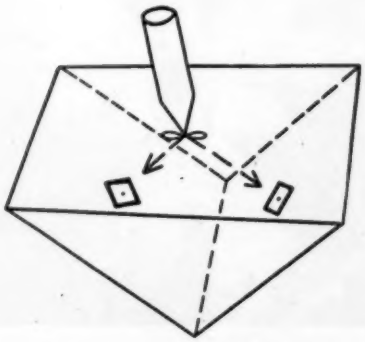

FIG. 14. Superficial and transmitted patterns formed by punching on a dodecahedron face.

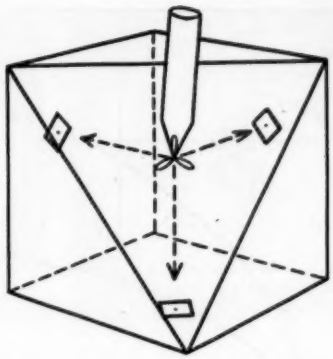

FIG. 15. Superficial and transmitted patterns formed by punching on an octahedron face.

cubic face the inner strains show a fourfold symmetry which coincides with the symmetry of the superficial patterns (Fig. 15). In the dodecahedron face (110) the inner strains in the wing direction are only slightly shorter than those perpendicular to it. The twofold symmetry is not so well-pronounced as that in the superficial wings. On the other hand, the inner strains on the octahedron face (111) show sharp threefold symmetry of the same quality as that on cubic face. All inner strains are unstable and disappear in a few months at room temperature, and more quickly at higher temperatures.

Although the general process of plastic deformation of silver chloride crystals[16] is the same as that of thallium halides, they give only diffuse patterns under the pressure of a cone, as shown in Fig. 16. This different behavior may be a result of inhomogeneous deformation caused by the high plasticity of silver chloride.

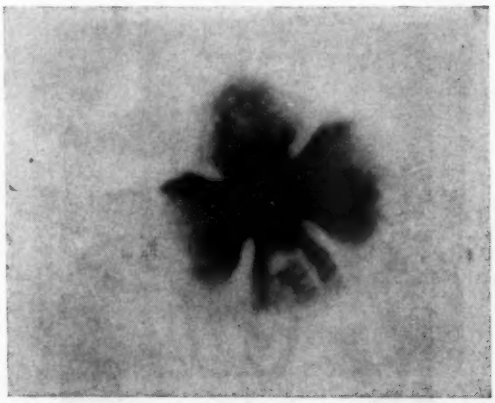

FIG. 16. Punch pattern on silver chloride crystal.

[16] J. F. Nye, Nature 162, 300 (1948).

FIG. 17. Punch pattern on cubic face of NaCl.

Sodium chloride has already been investigated by various authors using a punch needle together with microscopic observation.[17] The patterns obtained by punching a cone coincide completely with previous results. In both cases, the applied force causes splitting along the dodecahedron faces (Fig. 17). The four wings are not always symmetrical; a deviation of some degrees from the dodecahedron face can be observed. Sometimes only two or three wings appear instead of four. The splitting is caused by low plasticity of sodium chloride crystals.

[17] Vraski, Gogoberidze, and Flerova, J. Tech. Phys. **17**, 723 (1947).

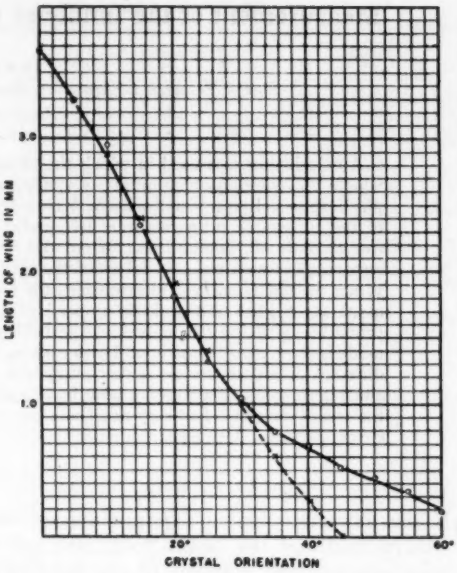

FIG. 18. Length of superficial wings in dependence of crystal orientation. 0—0 experimental values; ×—× theoretical values.

At the temperature of 400 degrees centigrade the punch patterns of sodium chloride crystals are similar to those of silver chloride crystals at room temperature.

Measurement of the Influence of Local Adaptation on Color Matching*

DAVID L. MACADAM
Kodak Research Laboratories, Eastman Kodak Company, Rochester, New York
(Received March 14, 1949)

A colorimeter was used in which the adjustable primaries can be extinguished separately, and in which the comparison stimulus is synthesized by an equivalent optical system, thus assuring a physical as well as a visual match. After a few matches were obtained in the customary manner, two of the adjustable primaries were left undisturbed and the third was merely flashed intermittently, five seconds off, one-half second on. After observation of one or two cycles, a great difference is apparent between the two halves of the field during the half-second when all three primaries are on. When the red primary is intermittent, the lower semicircle appears bluish-green for five seconds, and much too reddish to match the upper half during the brief flash. The observers then readjusted the primaries until the two halves of the field appeared to match during the flash. Each primary was made intermittent, in turn, and also each pair of primaries, giving measurements of the effect for six different directions around each comparison color. Achromatic and moderately saturated red, green, and purple comparison colors were investigated. The intermittent stimulus which appears to match the steady stimulus lies very nearly on the straight line in the chromaticity diagram connecting the latter with the adapting stimulus.

ON one occasion during experiments on color discrimination,[1] the beam from one of the primaries illuminating one-half of the field of an additive colorimeter was obstructed for a few seconds, and then restored. The observer happened to keep his gaze fixed on the colorimeter field and noticed that the formerly well-matched halves of the field appeared strikingly different during the first few seconds after the beam was restored. After a few seconds further observation, the appearance of mismatch vanished, rather rapidly, but not abruptly. Needless to say, the half of the field in which the primary was obstructed had a very different color from the undisturbed half, during the period of obstruction. Steady fixation and observation of the field during that time apparently adapted the corresponding portion of the retina differently from the portion exposed to the normal stimulus of the undisturbed half of the field. That this adaptation was local and confined to a sharply defined area of the retina was verified by shifting the gaze at the moment that the obstructed primary was restored, when a patch of the same shape as the disturbed field and the same chromatic difference as previously noticed was seen superimposed on the complete field but displaced corresponding to the shift of gaze. This patch could be moved at will during the few seconds that it persisted by directing the center of gaze to various parts of the colorimeter field.

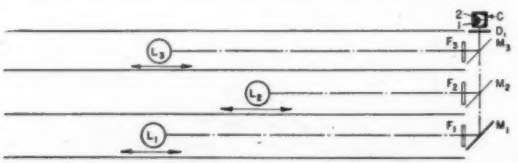

FIG. 1. Elevation, showing arrangement of adjustment section of additive colorimeter. Diagonal, semitransparent mirrors reflect and transmit beams from lamps into photometric cube, *C*. Primary filters are shown to left of mirrors.

Such sharply defined patches of apparent color difference, following exposure to contrasting stimuli, are often called "after-images," but it seems more helpful to an understanding of the phenomenon to describe this effect as a consequence of local adaptation. Its kinship to adaptation, as generally understood, seems evident, both from the conditions under which it arises and from the similarity of effects.

The arrangement of the colorimeter used to measure the influence of adaptation is shown in Figs. 1–3. The front view, shown in Fig. 1, shows the movable lamps, L_1, L_2, and L_3, the primary filters, F_1 (red), F_2 (green), and F_3 (blue), the opaque mirror, M_1, and the semitransparent mirrors, M_2 and M_3, which reflect the three primary beams onto the diffuser, D_1, which is viewed through a photometric cube, *C*. That diffuser, which is illuminated by the adjustable primaries, is viewed in the lower half, (1), of the circular field (subtending 2 degrees) indicated on the face of *C*. The upper half (2) of the photometric field is filled with light from the second diffuser, D_2, which is seen in the end view of the instrument (Fig. 2), together with the mirrors, M_4 to M_7, and the lamp, L_4,

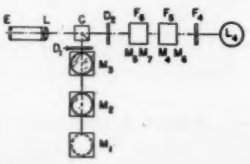

FIG. 2. End view, showing relative placement of adjustment and comparison systems of colorimeter.

* Communication No. 1248 from the Kodak Research Laboratories.

[1] W. R. J. Brown and D. L. MacAdam, J. Opt. Soc. Am. 38, 1093 (1948).

TABLE I. Colorimetric specifications of local adaptation and matching stimuli.

Adaptation	Observer	Components B	G	R	Match g	r	Adaptation x	y	Y (ft. L)	Match x	y	Y (ft. L)
B+G+R (steady)	WRJB	1.760	1.108	0.174	0.364	0.057	0.336	0.300	3.88	0.336	0.300	3.88
	DLM	1.715	1.120	0.176	0.372	0.058	0.336	0.304	3.90	0.336	0.304	3.90
	HRC	1.740	1.098	0.188	0.363	0.062	0.342	0.301	3.45	0.342	0.301	3.45
G+R (intermittent)	WRJB	0.730	1.215	0.154	0.578	0.073	0.475	0.491	3.36	0.393	0.391	3.61
	DLM	0.490	1.320	0.132	0.677	0.068	0.454	0.508	3.42	0.396	0.430	3.59
	HRC	0.475	1.232	0.177	0.654	0.094	0.486	0.484	3.53	0.432	0.414	3.70
B+R	WRJB	1.972	0.751	0.212	0.256	0.072	0.350	0.169	2.01	0.344	0.259	3.49
	DLM	2.135	0.597	0.220	0.202	0.075	0.346	0.167	2.12	0.342	0.242	3.29
	HRC	0.976	0.149	0.166	0.115	0.129	0.395	0.186	1.38	0.378	0.222	1.67
B+G	WRJB	1.905	1.385	0.094	0.409	0.028	0.223	0.319	3.38	0.287	0.317	3.97
	DLM	2.040	1.352	0.058	0.392	0.017	0.220	0.306	3.36	0.258	0.306	3.73
	HRC	1.760	1.175	0.052	0.393	0.017	0.220	0.306	2.92	0.260	0.306	3.24
R	WRJB	1.220	0.902	0.200	0.388	0.085	0.705	0.306	1.26	0.383	0.315	3.45
	DLM	0.921	0.758	0.207	0.402	0.110	0.705	0.306	1.30	0.418	0.325	3.11
	HRC	0.544	0.292	0.204	0.281	0.196	0.705	0.306	1.28	0.488	0.293	2.04
B	WRJB	2.055	0.936	0.110	0.302	0.035	0.155	0.094	0.71	0.286	0.269	3.24
	DLM	1.945	0.830	0.056	0.293	0.020	0.155	0.094	0.67	0.251	0.258	2.65
	HRC	2.560	0.831	0.057	0.241	0.016	0.155	0.094	0.88	0.230	0.227	2.87
G	WRJB	0.866	1.362	0.094	0.587	0.040	0.316	0.620	2.68	0.341	0.395	3.57
	DLM	1.128	1.368	0.090	0.529	0.035	0.316	0.620	2.69	0.326	0.371	3.65
	HRC	0.619	1.104	0.057	0.621	0.032	0.316	0.620	2.17	0.330	0.415	2.80

which illuminate it. The arrangement of these mirrors which provide the fixed comparison stimulus, is shown in Fig. 3, which is a top view of the instrument. Light from the single lamp, L_4, is transmitted by the red filter, F_4, and by the semitransparent mirrors, M_4 and M_5, to the diffuser, D_2, which is viewed through the upper half (2) of the photometric cube. Another beam from the lamp, L_4, is reflected from the opaque mirror, M_6, through the green filter, F_5, and by the semitransparent mirror, M_4. The third beam from L_4 is reflected from the opaque mirror, M_7, through the blue filter, F_6, and by the semitransparent mirror, M_5, to the diffuser, D_2. The face of the photometric cube, C, is examined through the lens, L, monocularly by the observer's eye at E.

The settings of the adjustable lamps, L_1, L_2, L_3, required to match the fields with steady illumination of both were first determined. Then, with one or two lamps flashing intermittently, 5 seconds off, one-half second on, the positions of the lamps were readjusted so that the two halves of the field appeared to match during the half-second periods when all three lamps were lit. In general, the positions of all three lamps require readjustment, but the major changes are of the intermittently flashing lamps, in directions such as to reduce their contributions to the part of the field in which the adaptation is taking place.

While the intermittently flashing lamps are extinguished, the corresponding part of the field is illuminated by the remaining primaries, resulting in highly chromatic adaptation stimuli. Thus, when the lamp furnishing the red primary is extinguished, the lower half of the field appears as a saturated bluish-green, of less brightness than the steady top half of the field because of the absence of the red component. Before the positions of the lamps are readjusted, while they remain in the positions required for a steady-state match, the momentary restoration of the red primary makes the lower half of the field appear too red and too bright. After the positions of the lamps are adjusted to yield an apparent match during the momentary restorations of the red primary, their contributions are such as to make the lower half of the field appear distinctly bluish-green if compared to the top in steady-state observation.

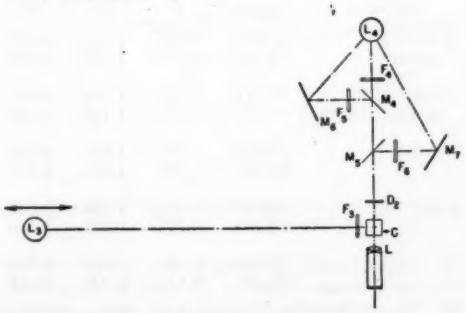

FIG. 3. Plan showing arrangement of lamp, L_4, mirrors, and filters in comparison stimulus system. Beam passing directly through provides red primary. Beam reflected from left passes through green filter. Top of one movable lamp, L_3, and the photometric cube, C, and eyepiece, E, are shown.

TABLE II. Colorimetric specifications of local adaptation and matching stimuli.

Adaptation	Observer	Components B	G	R	Match g	r	Adaptation x	y	Y (ft. L)	Match x	y	Y (ft. L)
B+G+R (steady)	WRJB	2.422	0.286	0.189	0.098	0.065	0.306	0.186	2.58	0.306	0.186	2.58
	DLM	2.280	0.290	0.171	0.106	0.062	0.305	0.194	2.43	0.305	0.194	2.43
	HRC	2.245	0.240	0.173	0.090	0.065	0.307	0.187	2.33	0.307	0.187	2.33
G+R (intermittent)	WRJB	1.862	0.330	0.181	0.139	0.076	0.596	0.395	1.78	0.331	0.211	2.43
	DLM	1.540	0.291	0.175	0.145	0.087	0.625	0.372	1.67	0.350	0.218	2.20
	HRC	0.857	0.305	0.186	0.226	0.138	0.580	0.408	1.77	0.406	0.267	2.07
B+R	WRJB	2.473	0.194	0.178	0.068	0.063	0.300	0.148	1.97	0.301	0.173	2.35
	DLM	3.835	0.160	0.205	0.038	0.049	0.270	0.137	2.61	0.272	0.151	2.92
	HRC	2.240	0.110	0.157	0.044	0.063	0.302	0.148	1.76	0.302	0.164	1.98
B+G	WRJB	2.928	0.363	0.128	0.106	0.037	0.177	0.157	1.72	0.260	0.181	2.53
	DLM	2.762	0.307	0.123	0.096	0.039	0.174	0.147	1.56	0.260	0.172	2.33
	HRC	2.048	0.300	0.052	0.125	0.022	0.179	0.163	1.30	0.237	0.178	1.62
R	WRJB	1.779	0.317	0.187	0.139	0.082	0.705	0.306	1.17	0.340	0.208	2.41
	DLM	2.465	0.252	0.168	0.087	0.058	0.705	0.306	1.06	0.293	0.180	2.40
	HRC	0.890	0.154	0.212	0.123	0.169	0.705	0.306	1.33	0.422	0.240	1.94
B	WRJB	3.082	0.356	0.138	0.100	0.039	0.155	0.094	1.06	0.264	0.175	2.63
	DLM	3.205	0.198	0.126	0.056	0.039	0.155	0.094	1.10	0.259	0.158	2.29
	HRC	2.085	0.140	0.058	0.061	0.025	0.155	0.094	0.72	0.228	0.145	1.36
G	WRJB	2.505	0.341	0.166	0.113	0.055	0.316	0.620	0.67	0.296	0.193	2.58
	DLM	2.178	0.289	0.144	0.111	0.055	0.316	0.620	0.57	0.292	0.192	2.22
	HRC	1.754	0.205	0.097	0.100	0.047	0.316	0.620	0.43	0.276	0.180	1.62

The colorimetric specifications of the readjusted field can be computed from the amounts of the primaries, and constitute the measure of the effect of the local adaptation. The adapting stimulus can be computed from the amounts of the one or two primaries which are not extinguished during each set of observations. These amounts are changed by the process of readjusting the lower field to produce a momentary match. Consequently, the adapting stimulus is rarely the same, even in successive repetitions of the same experiment, but each adapting stimulus is determined and is specified by the recorded amounts of the readjusted primaries.

The amounts by which the various primaries have to be changed to obtain a match during the momentary combination of all three primaries are indicated in Tables I to IV for four different choices of the steady stimulus with six different adapting stimuli surrounding each choice. Two observers each made all of these observations, and repeated their determinations on a second day for the last of the steady stimuli (Table IV). A third observer made observations for the first two of the steady stimuli (Tables I and II). The first line of data in each table for each observer indicates the steady stimulus and the amounts of the primaries

TABLE III. Colorimetric specifications of local adaptation and matching stimuli.

Adaptation	Observer	Components B	G	R	Match g	r	Adaptation x	y	Y (ft. L)	Match x	y	Y (ft. L)
B+G+R (steady)	WRJB	0.268	1.145	0.053	0.782	0.036	0.361	0.485	2.68	0.360	0.485	2.68
	DLM	0.239	1.135	0.050	0.797	0.035	0.360	0.489	2.63	0.361	0.489	2.63
G+R (intermittent)	WRJB	0.084	1.194	0.051	0.900	0.038	0.397	0.558	2.67	0.382	0.530	2.70
	DLM	0.152	1.120	0.050	0.847	0.038	0.400	0.556	2.52	0.374	0.510	2.57
B+R	WRJB	0.374	1.070	0.084	0.701	0.055	0.463	0.215	0.66	0.382	0.444	2.76
	DLM	0.205	1.032	0.059	0.797	0.046	0.497	0.227	0.44	0.380	0.485	2.47
B+G	WRJB	0.273	1.200	0.033	0.798	0.022	0.287	0.525	2.46	0.334	0.500	2.66
	DLM	0.205	1.195	0.036	0.832	0.025	0.294	0.550	2.42	0.344	0.516	2.65
R	WRJB	0.303	1.061	0.085	0.733	0.059	0.705	0.306	0.53	0.388	0.455	2.73
	DLM	0.152	0.906	0.061	0.811	0.055	0.705	0.306	0.38	0.370	0.498	2.22
B	WRJB	0.345	1.208	0.052	0.753	0.032	0.155	0.094	0.12	0.347	0.474	2.82
	DLM	0.178	0.770	0.034	0.784	0.035	0.155	0.094	0.06	0.343	0.493	1.79
G	WRJB	0.160	1.220	0.038	0.860	0.027	0.316	0.620	2.40	0.348	0.526	2.69
	DLM	0.122	1.292	0.028	0.897	0.019	0.316	0.620	2.55	0.338	0.547	2.76

TABLE IV. Colorimetric specifications of local adaptation and matching stimuli.

Adaptation	Observer	Components B	G	R	Match g	r	Adaptation x	y	Y (ft. L)	Match x	y	Y (ft. L)
B+G+R (steady)	WRJB(1)	1.542	1.115	1.620	0.261	0.379	0.596	0.309	12.90	0.596	0.309	12.90
	WRJB(2)	1.660	1.375	1.900	0.289	0.385	0.597	0.312	15.20	0.597	0.312	15.20
	DLM(1)	1.765	1.312	1.908	0.263	0.383	0.593	0.309	15.17	0.593	0.309	15.17
	DLM(2)	1.174	1.002	1.355	0.284	0.384	0.594	0.312	10.88	0.594	0.312	10.88
G+R (intermittent)	WRJB(1)	0.442	1.352	1.600	0.399	0.472	0.656	0.347	12.70	0.640	0.338	12.85
	WRJB(2)	0.395	1.665	1.808	0.430	0.478	0.652	0.350	14.62	0.639	0.343	14.75
	DLM(1)	0.449	1.343	1.809	0.373	0.503	0.661	0.343	13.98	0.647	0.336	14.14
	DLM(2)	0.360	0.820	1.421	0.315	0.546	0.667	0.337	10.54	0.647	0.329	10.67
B+R	WRJB(1)	1.510	0.417	1.725	0.114	0.472	0.633	0.278	10.34	0.620	0.293	12.16
	WRJB(2)	1.752	0.405	1.985	0.098	0.480	0.629	0.276	13.06	0.619	0.288	13.86
	DLM(1)	2.040	0.681	1.945	0.146	0.417	0.617	0.271	12.90	0.599	0.291	14.24
	DLM(2)	1.708	0.550	1.475	0.147	0.395	0.606	0.267	9.85	0.589	0.290	10.93
B+G	WRJB(1)	1.455	1.630	0.656	0.436	0.175	0.249	0.394	3.71	0.494	0.346	7.93
	WRJB(2)	2.348	2.830	1.108	0.450	0.176	0.245	0.380	6.38	0.494	0.340	13.39
	DLM(1)	1.545	1.955	0.682	0.468	0.163	0.246	0.383	4.38	0.481	0.344	8.66
	DLM(2)	1.360	1.480	0.645	0.425	0.185	0.243	0.375	3.38	0.498	0.336	7.45
R	WRJB(1)	0.521	0.780	1.745	0.256	0.573	0.705	0.306	10.94	0.643	0.317	12.65
	WRJB(2)	0.678	0.811	2.015	0.231	0.575	0.705	0.306	12.64	0.643	0.314	14.47
	DLM(1)	0.670	0.750	1.984	0.202	0.583	0.705	0.306	12.45	0.653	0.315	14.16
	DLM(2)	0.491	0.516	1.364	0.218	0.576	0.705	0.306	8.56	0.640	0.313	9.74
B	WRJB(1)	1.660	1.485	1.060	0.352	0.252	0.155	0.094	0.57	0.536	0.319	10.16
	WRJB(2)	2.200	2.023	1.785	0.337	0.297	0.155	0.094	0.76	0.560	0.319	15.94
	DLM(1)	2.228	1.560	0.690	0.348	0.154	0.155	0.094	0.77	0.479	0.312	8.17
	DLM(2)	1.212	1.060	0.600	0.369	0.209	0.155	0.094	0.42	0.506	0.319	6.27
G	WRJB(1)	0.810	1.990	1.055	0.503	0.274	0.316	0.620	3.92	0.565	0.357	10.82
	WRJB(2)	1.485	2.650	1.430	0.475	0.257	0.316	0.620	5.22	0.549	0.349	14.70
	DLM(1)	0.976	1.940	0.590	0.553	0.168	0.316	0.620	3.82	0.501	0.368	7.86
	DLM(2)	0.711	1.785	0.557	0.585	0.182	0.316	0.620	3.51	0.520	0.376	7.25

necessary for the steady-state match. Each subsequent line of data indicates the kind of adaptation, the modified amounts of the primaries required for a momentary match, and the colorimetric specification of the stimulus required for the momentary match under each condition of adaptation indicated.

The connections between the amounts of the primaries used and the conventional colorimetric specifications are given by the following calibration equations:

$$x=X/(X+Y+Z), \quad y=Y/(X+Y+Z),$$

where

$$X=0.577B+1.026G+14.37R.$$
$$Y=0.345B+1.967G+\ 6.27R.$$
$$Z=2.887B+0.080G+\ 0.20R.$$

The calibrations are slightly different for the three observers, but the differences are negligible in comparison with the magnitudes of the effects of adaptation and the relatively poor repeatability of the observations, as indicated in Table IV.

The results are shown on chromaticity diagrams in Figs. 4 to 7, in which it appears that most of the stimuli required for a momentary match are close to the straight lines connecting the steady comparison stimulus and the adapting stimulus. Some

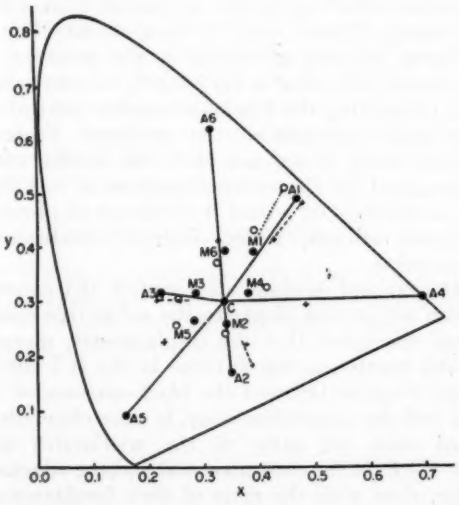

FIG. 4. Effect of adaptations on momentary matches for nearly neutral stimulus, C. A_1 to A_6 represent six different adaptations. Solid dots indicate adaptations for WRJB, open circles for DLM, crosses for HRC. M_1 indicates chromaticities necessary to match C after adaptation to A_1, with same code for the three observers. Straight lines from C toward adaptation chromaticities pass near chromaticities required for match.

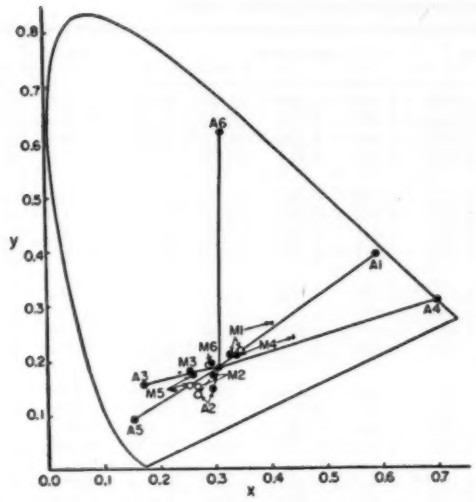

FIG. 5. Effects of adaptations on momentary matches for purple stimulus. Same code as in Fig. 4.

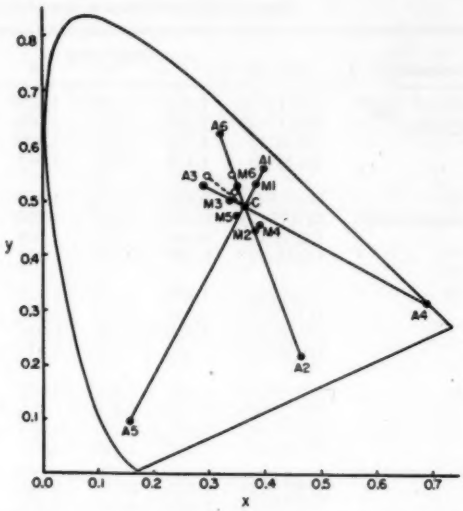

FIG. 6. Effects of adaptations on momentary matches for green stimulus, *C*. Same code as in Figs. 4 and 5 for observers WRJB and DLM only.

of the sets of points differ consistently, and perhaps significantly, from those lines. A few results specified in Tables II and III are so inconsistent with the general pattern, and with the subjective effects of the adaptations, that they are not shown in Figs. 5 and 6. The analogy of the results shown in Figs. 4 to 7 with the results reported by Helson and Michels[3] is suggestive. After determining the stimulus which is chosen as neutral with a black surround, Helson and Michels found that the stimulus selected as neutral in the presence of a chromatic surround is very nearly on the straight line connecting the black-surrounded neutral with the point representing the surround. From the present work it appears that this finding can be generalized for chromatic stimuli so as to indicate the conditions for visual equivalence of physically different stimuli, under different conditions of adaptation.

Helson and Michels also studied the extent to which adaptation displaces the subjective neutral. They concluded that the displacement, measured as the fraction of the distance in the ICI chromaticity diagram between the black-surrounded neutral and the adaptation color, is more closely correlated with the ratio of the tristimulus sums, $X+Y+Z$, of the chosen neutral and the adaptation color, than with the ratio of their luminances, Y. The same conclusion has been drawn from a study of the present data, although the scatter of the points representing the results seems greater than would be expected from the reproducibility of the determinations. The direct-matching technique used in the present work, even though the observers were handicapped by only momentary comparisons, yields precision and encourages confidence in the adjustments that does not seem to be adequately represented by the scatter of the points in Fig. 8. The abscissa, R, is the ratio of the tristimulus sum, $X+Y+Z$, of the momentarily matched stimulus to the tristimulus sum of the adapting stimulus. The ordinate, r, is the fraction of the distance from the steady stimulus to the

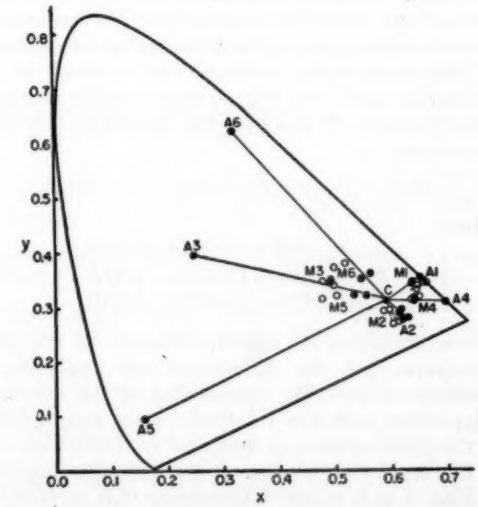

FIG. 7. Effects of adaptation on momentary matches for red stimulus, *C*. Same code as in Figs. 4, 5, and 6. Two results shown for each observer and condition, indicating repeatability.

[3] H. Helson and W. S. Michels, J. Opt. Soc. Am. **38**, 1025–1033 (1948).

adapting stimulus marked off by the stimulus required for momentary match. The scatter of the points is perhaps not surprising, since the ICI tristimulus values and their sums, used here as by Helson and Michels, are highly arbitrary in nature and according to present indications not closely related to the physiological responses underlying color perception. The superiority of the ratio of the tristimulus sums over the ratio of luminances may indicate the most profitable way to treat adaptation data when the physiological primaries are determined. If so, then the somewhat disappointing correlation with ICI tristimulus values can be understood. Some years ago, the author demonstrated that if adaptation phenomena can be accounted for by a simple formula in terms of some set of primaries, no other significantly different set of primaries can be successfully substituted.[3] In the present instance, we do not know what set of primaries would correlate with our results within the precision of our data, but cannot be surprised that the arbitrarily chosen ICI primaries do not.

Helson and Michels found that the formula,

$$r=1/(1+0.21R^2),$$

gave the closest fit to their results. This function is shown by the dotted curve in Fig. 8. In general, it is too high to fit our results. The broken-line curve shows the result of merely changing the coefficient in their formula,

$$r=1/(1+0.54R^2).$$

This gives a closer fit to the present results, but increasing the exponent of R seems to be even more successful,

$$r=1/(1+0.54R^3).$$

This function is shown in Fig. 8 by the solid curve. Perhaps an exponent between 2 and 3 would yield a better fit to the results, but the scatter of the points hardly justifies such elaboration.

In conclusion, the empirical rule formulated by Helson and Michels seems to be confirmed and generalized qualitatively, and modified only slightly quantitatively to an extent that may indicate the effect of the radical difference in adaptation and observing conditions. After adaptation, the chromaticity needed to match a self-adapting stimulus lies close to the straight line between the points representing the self-adapting stimulus and the contrasting adaptation color. The distance between the points representing the stimuli matched under those conditions, expressed as a fraction of the distance between the points representing the adapting stimuli, correlates with the ratio of the tristimulus sums of the matching stimulus and its adaptation. This correlation is somewhat better than if the ratio of the luminances is used.

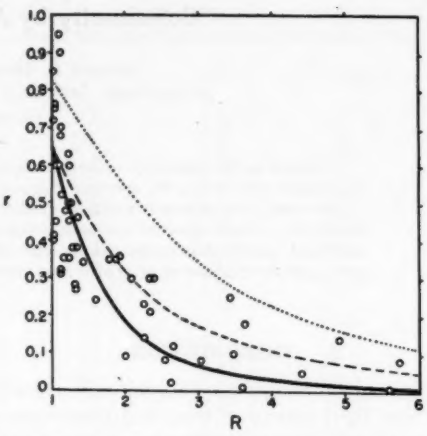

FIG. 8. Correlation between effects of adaptations and magnitudes of adapting stimuli. Abscissas are ratio of ICI tristimulus sums: that of the matching stimulus divided by that of the adapting stimulus. Ordinates are distances in ICI diagram between matched stimuli expressed as fractions of distance from self-adapting comparison stimulus to adapting stimulus of matching field.

ACKNOWLEDGMENT

Our thanks are tendered here to Walter R. J. Brown, Harold R. Condit, and Robert W. Burnham, for making observations which were of a particularly tantalizing kind.

[3] R. M. Evans, J. Opt. Soc. Am. **33**, 579–614 (see pp. 600, 601) (1943).

Colorimetry by Abridged Spectrophotometry

ARTHUR C. HARDY AND F. MANSFIELD YOUNG
Massachusetts Institute of Technology, Cambridge 39, Massachusetts
(Received March 4, 1949)

It seems to be generally conceded that the results obtained with a photoelectric colorimeter are significant only when the spectral characteristics of the sample and standard are similar. When this is the case, the selected ordinate method of integrating spectrophotometric data can be greatly abridged, a single selected ordinate being adequate under many circumstances. This means that an abridged spectrophotometer can be used for the dual purpose of production control and, when necessary, the determination of tristimulus values.

INTRODUCTION

IT is easy to imagine a photoelectric colorimeter whose light source, filters, and phototube meet the requirements imposed by the nature of the task that such an instrument is expected to perform. In practice, it is quite another matter to meet these requirements. Much effort has been expended in the attempt to do so during the quarter century that phototubes have been commercially available; but the most that can be claimed for a photoelectric colorimeter is the ability to measure a small difference in color between a sample and a standard when their spectrophotometric characteristics are not too different.

Small color differences of this sort occur most frequently in manufacturing processes where the color of the finished articles must not be allowed to vary significantly from the color that has been adopted as standard. For the control of such processes, a spectrophotometer is inherently better suited than a colorimeter because spectrophotometric measurements are easier to interpret in terms of the behavior of dyes and pigments. When a spectrophotometer is used for this purpose, the usual procedure can be considerably abridged; and, with the development of suitable interference filters to replace the customary prismatic dispersing system, an abridged spectrophotometer can be an exceedingly simple instrument. Its filters need not be fitted to the characteristics of the light source and phototube like those used in a photoelectric colorimeter. Nor is there need to be concerned over the effect of changes in the spectral quality of the light from the source or in the spectral sensitivity of the phototube.

The purpose of this paper is to show how, when the color difference is one that can be measured with a photoelectric colorimeter, the evaluation of the color difference by spectrophotometry and integration can be greatly simplified. In the limit, a comparison of sample and standard at three properly selected wave-lengths will often suffice for the determination of the three tristimulus values.

SPECIAL SETS OF SELECTED ORDINATES

Let R be a function of wave-length representing the spectral reflectance of a color standard with which samples from a production line are to be compared. In terms of the symbols that have been widely adopted in the field of colorimetry,[1] the tristimulus values of the standard are

$$X=\int_0^\infty R\bar{E}\bar{x}d\lambda,$$

$$Y=\int_0^\infty R\bar{E}\bar{y}d\lambda,$$

$$Z=\int_0^\infty R\bar{E}\bar{z}d\lambda.$$

For the purpose of this paper, it will be sufficient to consider only calculations involving X, inasmuch as the same principles apply to calculations involving Y and Z.

The usual method of integration is to approximate the integral of $R\bar{E}\bar{x}$ by the finite sum

$$X=\sum_0^\infty R\bar{E}\bar{x}\Delta\lambda.$$

Then, using published tables of selected ordinates that divide $\sum\bar{E}\bar{x}\Delta\lambda$ into equal parts, the numerical value of the integral is found by summing the values of R at the selected wave-lengths, dividing the sum by the number of entries, and multiplying by the proper normalizing factor. For a given accuracy in the result, the number of ordinates that must be used depends on the complexity of the function R. If the standard happens to be white, a single ordinate will often suffice.

Now let $\epsilon=R'/R$, where R' is the spectral reflectance of a sample that is intended to be a color match for the standard. Making the same approxi-

[1] *Handbook of Colorimetry* (The Technology Press, Cambridge, 1936).

mation as before

$$X' = \int_0^\infty R' E\bar{x} d\lambda,$$

$$= \sum_0^\infty R' E\bar{x} \Delta\lambda,$$

$$= \sum_0^\infty \epsilon R E\bar{x} \Delta\lambda.$$

From the procedure outlined in the paragraph above, it will be evident that $\sum RE\bar{x}\Delta\lambda$ can be divided into equal parts to determine a set of selected ordinates that bear the same relationship to ϵ that the standard set bears to R. From the nature of the circumstances, the value of ϵ at any wave-length is ordinarily in the neighborhood of unity. A single ordinate may therefore suffice for the integration.

AN ILLUSTRATIVE EXAMPLE

For the case of a yellow surface, adopted here as the standard color for the purpose of this illustration, column two in Table I lists measured values of R at the wave-lengths appearing in column one. Column three lists values of $E_c\bar{x}$ from Table XV of the Handbook of Colorimetry.[1] The fourth column lists the running sums of $RE_c\bar{x}$. Then, the sum of the entries in column three being 1042.7, and the normalizing factor 0.9804,

$$X = 63{,}319/1042.7(0.9804) = 59.54.$$

This is essentially the weighted ordinate method of integration.

If a single selected ordinate is desired for the integration of other colors in the neighborhood of this standard, the wave-length of the ordinate can be found by dividing 63,319 by 2 and interpolating in column four. By inspection, the running sum $63{,}319/2 = 31{,}660$ is seen to occur in the wave-length interval 595–605 in column five; and a linear interpolation within this interval shows the wave-length of this ordinate to be 596 mμ.

For the purpose of this illustration, another yellow surface was prepared whose spectrophotometric curve R' was of the same general shape as R. Using an automatic integrator employing 98 selected ordinates, X' was found to be 56.54. The single ordinate method of integration was then tried by measuring $\epsilon = R'/R$ at 596 mμ, the result being $\epsilon = 0.9475$. Hence,

$$X' = (0.9475)X = (0.9475)(59.54) = 56.41.$$

This value is in reasonably good agreement with the value of 56.54 obtained by the 98 ordinate method.

A plot of the function $RE_c\bar{x}$ from the data in the above table shows that the function has two maxima, one at 440 mμ and the other at 600 mμ. When this is the case, it is sometimes advantageous to divide the integral into two parts and to integrate each part separately. That is, for the case of the standard yellow surface of this illustration, we may write that

$$\sum_0^\infty RE\bar{x}\Delta\lambda = \sum_0^{505} RE\bar{x}\Delta\lambda + \sum_{505}^\infty RE\bar{x}\Delta\lambda.$$

The numerical value of the first sum is shown by column four of the table to be 1539. Dividing this sum by 2 and interpolating as before, gives 449 mμ as the wave-length of the single selected ordinate for the first integration. The numerical value of the second sum is $63{,}319 - 1539 = 61{,}780$. Dividing this sum by 2 and adding 1539 gives 32,429, which is seen by interpolation to correspond to 597 mμ, which is the wave-length of the selected ordinate for the second integration. The measured values of ϵ at these wave-lengths were found to be 1.0362 and 0.9474, respectively. These values must be weighted by the factors 1539/63,319 and 61,780/63,319. Hence,

$$X' = \left[\frac{(1539)(1.0362)}{(63{,}319)} + \frac{(61{,}780)(0.9474)}{(63{,}319)}\right] 59.54,$$

$$= 56.54.$$

TABLE I.

λ	R	$E\bar{x}$	$\sum_0^{\lambda+5} RE\bar{x}\Delta\lambda$	$\lambda+5$
400	7.4	0.91	7	405
410	7.6	3.48	33	415
420	7.6	13.19	133	425
430	7.5	31.92	373	435
440	7.3	42.15	681	445
450	7.2	41.69	981	455
460	7.1	35.81	1235	465
470	7.3	24.19	1412	475
480	7.7	11.85	1503	485
490	8.2	3.86	1534	495
500	9.1	0.55	1539	505
510	13.0	0.95	1552	515
520	23.3	6.13	1695	525
530	40.5	16.22	2352	535
540	57.7	29.65	4062	545
550	66.8	45.60	7108	555
560	71.0	62.60	11553	565
570	73.3	77.97	17268	575
580	74.8	89.61	23971	585
590	75.9	95.68	31233	595
600	76.8	95.26	38549	605
610	77.6	88.65	45428	615
620	78.4	75.28	51330	625
630	79.1	56.54	55803	635
640	79.9	39.32	58944	645
650	80.7	25.00	60962	655
660	81.5	14.50	62144	665
670	82.4	7.54	62765	675
680	83.3	3.93	63092	685
690	84.1	1.82	63245	695
700	84.9	0.87	63319	705

This is exactly the value that was obtained with the automatic integrator using 98 ordinates.

SPECIAL METHODS OF INTEGRATION

When performing an integration by either the selected ordinate method or the weighted ordinate method, the number or ordinates should always be so large that, if a larger number were used, the results would not be altered significantly. It is well known that, if R is substantially constant throughout the spectrum, a single ordinate will suffice for the integration of $R\bar{x}$; and we have shown above that a single ordinate will suffice for the integration of $\epsilon R\bar{x}$ when ϵ is substantially constant throughout the spectrum and the integral of $R\bar{x}$ is already known.

Whenever the nature of the difference between sample and standard is known, it is always possible to devise a method, either theoretically or empirically, that will yield tristimulus values with a minimum number of isolated spectrophotometric measurements. Suppose, for example, that the difference in the spectrophotometric characteristics of sample and standard is due entirely to variations in the concentration of a substance whose behavior can be described by Beer's law. Let T^c represent the spectral transmittance of the standard and $T^{c+\Delta c}$ that of the sample. In this case,

$$X=\int_0^\infty T^c E\bar{x}d\lambda \quad \text{and} \quad X'=\int_0^\infty T^{c+\Delta c}E\bar{x}d\lambda.$$

But, expanding $T^{\Delta c}$ in a Taylor series,

$$X'=\int_0^\infty [1+(\Delta c/c)\ln T^c+\cdots 1/n!(\Delta c/c)^n(\ln T^c)^n+\cdots]T^cE\bar{x}d\lambda,$$

$$=X_0+(\Delta c/c)X_1+\cdots(\Delta c/c)^nX_n+\cdots,$$

where

$$X_n=\frac{1}{n!}\int_0^\infty (\ln T^c)^n T^c E\bar{x}d\lambda.$$

As a first approximation, consider only the first two terms of the expansion, which give

$$X'=X+\Delta c/c\int_0^\infty (\ln T^c)T^cE\bar{x}d\lambda.$$

It will be noticed that the integral of $(\ln T^c)T^cE\bar{x}$ does not involve the characteristics of the sample. Hence, when the numerical value of this integral has been ascertained, the difference between X' and X is found immediately by multiplying the numerical value of the integral by $\Delta c/c$. For the determination of $\Delta c/c$, a single measurement at any wave-length will suffice, but greater precision is obtained by making the necessary measurement at the wave-length of an absorption band. Evidently in this special case, all three tristimulus values can be found from a single measurement.

Visibility on Cathode-Ray Tube Screens. The Effect of Size and Shape of Pip*

N. R. BARTLETT,** S. B. WILLIAMS,*** AND R. M. HANES****
The Johns Hopkins University, Baltimore, Maryland

(Received February 11, 1949)

The area and contour of a stimulus have been indicated experimentally as factors in its visibility. In order to evaluate the role of such factors in radar presentation, an experiment was conducted to determine the effect upon visibility of changes in the size and shape of a radar pip. A wide range of beam widths (which determine the angular dimension of the pip) and pulse lengths (which determine its radial dimension) was investigated at each of three test field brightnesses. Results show that visibility, on a decibel scale, is a linear function of the logarithm of beam width at each of the three brightness levels used. For the dim background visibility is also a linear function of the logarithm of pulse length. For the brighter backgrounds, on the other hand, visibility is shown to be a curvilinear function of the logarithm of pulse length. It is suggested that the reason for the difference between an increase in beam width and an increase in pulse length lies in the time factor. Doubling the beam width, for example, results in a greater increase in the time during which the pip is exposed to the eye than does a corresponding increase in the pulse length.

THIS paper is concerned with the visibility of a radar pip—the visible effect of a radar signal or echo—as affected by changes in its length and width. On the screen of a cathode-ray tube of the intensity-modulated type, the pip appears as a local bright spot in the form of an arc segment. Changes in the dimensions of the pip are effected by altering the duration of the signal pulse or the width of the beam or lobe of energy.[1] The former determines the pip's radial aspect; the latter, its angular. In an immediate sense, this investigation is concerned only with the electrical properties of signals, although broadly conceived, it is a straightforward problem in visual perception.

Considered as the latter, it ought to be discussed in the context of the many recent researches on area and contour of visual targets. However, we have hesitated to embark on detailed comparisons with these experiments for the simple reason that the unique properties of a radar pip, visually considered, render comparisons exceedingly difficult. In the first place, the radar pip is an arc segment rather than the customary rectangle or circle. If it be allowed that an arc is approximately as visible as a rectangle—as indeed the work of Lamar, Hecht, Shlaer, and Hendley[2] would suggest, provided the two perimeters are equal—then one could regard the data of Brown[3] and of Brown and Niven[4] on rectangles and slits as directly relevant. But this allowance remains an assumption. In the second place, but more important, is the fact that most experimenters have exposed their stimuli instantaneously, or nearly so, and have employed backgrounds of uniform light distribution. Neither of these conditions prevails in radar. On an intensity-modulated scope, the observer sees the pip as a fading patch of light on a fading background. Furthermore in the PPI (Plan Position Indicator) type of scope, a radial line of the screen brightens as a rotating electronic beam excites it to fluorescence and all but this radial line is seen as fading phosphorescent light. In most screens, the fading also involves a color shift. As a further complication, Sweet and Bartlett[5] have shown that a bright pip persists longer and hence is visible for a longer time than is a dim one. All of these considerations make hazardous any prediction of radar visibility

TABLE I. Visual angles and transit times for the various beam widths and pulse lengths.

Pulse length (microseconds)	Visual angle subtended	Time of transit of overlay on each spoke (microseconds)	Beam width (degrees)	Visual angle subtended	Time of transit from leading edge to trailing edge (milliseconds)
0.5	1′43″	0.5	1	7′31″	8.3
1	3′29″	1	2	14′58″	16.7
2	6′56″	2	3	22′29″	25.0
3	10′25″	3	4	30′0″	33.3
4	13′54″	4	5	37′26″	41.7
10	34′48″	10	10	74′52″	83.3
20	69′35″	20	30	224′56″	250.0
30	104′21″	30	60	449′22″	500.0

* This research was carried out under Contract N5-ori-166, Task Order I, between Special Devices Center, ONR, and The Johns Hopkins University. This is Report No. 166-I-57, Project Designation No. NR-784-001, under that contract.

** Presently at Hobart College, Geneva, New York.

*** Presently at the College of William and Mary, Williamsburg, Virginia.

**** The authors are indebted to Mrs. Beverly Richards and Mr. Hamilton Mowbray for serving as subjects and to the engineering staff of the Systems Research Laboratory for technical assistance.

[1] S. B. Williams and N. R. Bartlett, "Visibility on cathode-ray tube screens: Problems and methods," J. Psych. **25**, 401 (1948).

[2] Lamar, Hecht, Shlaer, and Hendley, "Size, shape and contrast in detection of targets by daylight vision. I. Data and analytical description," J. Opt. Soc. Am. **37**, 531 (1947).

[3] R. H. Brown, "Spatial effects in human visual resolution," J. Gen. Psychol. **35**, 77 (1946).

[4] R. H. Brown and J. I. Niven, "The relation between the foveal intensity threshold and the length of an illuminated slit," J. Exper. Psych. **34**, 464 (1944); J. I. Niven and R. H. Brown, "Visual resolution as a function of intensity and exposure time in the human fovea," J. Opt. Soc. Am. **34**, 738 (1944).

[5] A. L. Sweet and N. R. Bartlett, "Visibility on cathode-ray tube screens: Signals on a P7 screen seen at different intervals after excitation," J. Opt. Soc. Am. **38**, 329 (1948).

from the facts of non-radar experiments. Very roughly speaking, we could expect that larger pips would be more visible, and that for certain restricted circumstances, Ricco's law should hold as a first approximation. According to Ricco's law, area and brightness would be interchangeable for small pips on a dim screen. For larger pips, the predictions would be uncertain indeed.

Outside of the standard visual literature, there are two recent contributions from engineering laboratories which are relevant to the radar problem. Hopkinson[6] has reported data for one observer with optical equipment designed to simulate pips on a PPI screen. Although his work appears to have been comprehensive and careful, his report is noticeably lacking in desired details. "A wide range of size and shape of trace was covered in these experiments," he writes. "An interesting discovery was the fact that over the range of values investigated, the visibility appeared to depend more on the area of the trace than on its shape" (pp. 798–799). The only supporting evidence presented

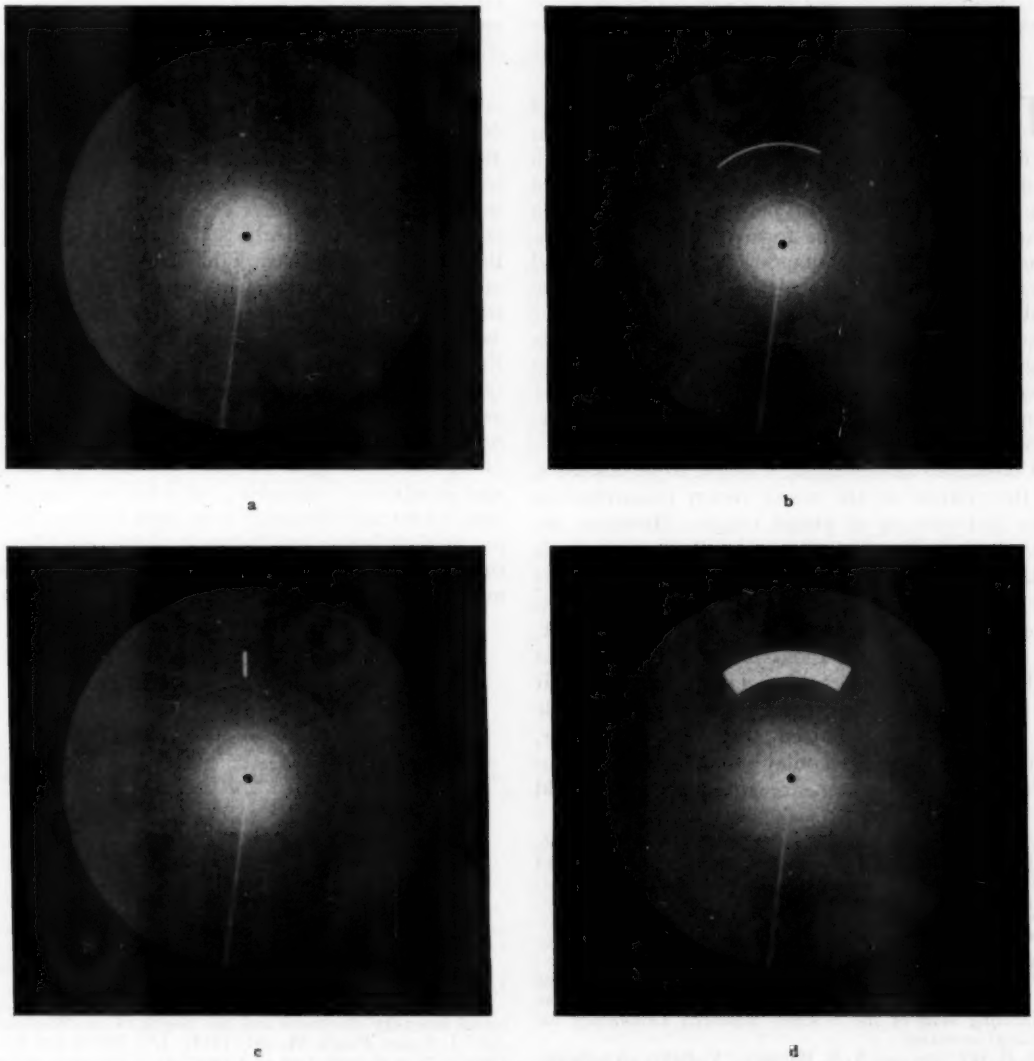

FIG. 1. Sample pips used in the experiment. a. Beam width = 1 degree; pulse length = $\frac{1}{2}$ microsecond. b. Beam width = 60 degrees; pulse length = $\frac{1}{2}$ microsecond. c. Beam width = 1 degree; pulse length = 30 microseconds. d. Beam width = 60 degrees; pulse length = 30 microseconds.

[6] R. G. Hopkinson, "Visibility of cathode-ray-tube traces in radar displays," J. Inst. Elec. Eng. 93, 795 (1946).

TABLE II. Visibility in decibels for the dim background.

Pulse length (microseconds)	Beam width (degrees) 1	2	3	4	5	10	30	60
0.5	11.5	12.5	13.5	14.0	14.5	14.5	17.0	18.0
1	13.0	14.0	15.0	16.0	15.5	17.5	18.0	20.5
2	14.5	15.5	16.5	18.5	19.0	19.0	20.0	24.5
3	15.5	17.5	18.5	19.5	18.5	20.5	23.0	23.0
4	15.5	18.0	19.0	19.5	22.0	22.5	26.0	27.0
10	19.0	20.0	23.5	23.5	24.0	25.0	27.5	30.0
20	20.5	21.5	23.5	25.5	25.5	28.5	30.5	29.5
30	20.5	20.5	25.0	24.5	24.5	26.0	29.0	31.0

TABLE III. Visibility in decibels for optimal background brightness.

Pulse length (microseconds)	Beam width (degrees) 1	2	3	4	5	10	30	60
0.5	23.0	24.5	25.0	23.5	27.5	28.0	30.0	32.0
1	29.5	30.0	31.5	29.5	33.0	33.0	36.0	40.0
2	33.5	34.5	35.0	37.5	37.0	39.0	41.0	44.0
3	35.0	36.5	39.0	38.0	38.5	39.0	41.5	47.5
4	36.0	37.0	39.5	41.5	40.5	43.0	44.5	42.5
10	36.5	36.5	40.5	39.0	40.0	42.0	44.0	48.0
20	37.5	39.0	39.5	41.0	42.0	43.5	44.0	47.5
30	36.5	40.5	40.5	40.5	40.5	42.5	45.0	46.0

is a series of monotonic but definitely curvilinear functions showing improvement, at several brightness levels, as area is increased. Nothing more is said about shape.

Payne-Scott[7] collected data on the effects of some electrical parameters determining pip size, using an artificial radar set with a real cathode-ray tube. Her data show, for example, that visibility, in decibels, increases as the $\frac{3}{4}$ power of the beam width over the limited range of 2° to 12° and that this relationship breaks down somewhere between 12° and 360°. Background brightness was not specified, making it difficult to interpret her data in visual theory. Having constructed a mathematical theory which failed to predict the $\frac{3}{4}$ power relationship even up to 12°, she remarks that the discrepancy between theory and fact may be due to, "the characteristic arc shape of the echo," which is "· · · of great aid in distinguishing it from noise pulses" (p. 187). (Noise pulses are pip-like appearances due to an assortment of electrical interferences.) She thereby tacitly recognizes shape as a perceptual factor of practical, engineering importance.

The present experiment can be regarded as an extension of the latter investigation. A standard radar viewing screen, of the PPI type, was employed and simulated echoes were the stimuli. Relevant properties of the signal at the final input stage, that is, the grid of the CR tube, were then related to their ultimate product: visibility. In our experiment the range of pip sizes and shapes was much greater than ever before studied or is likely to be encountered in operational radar. In addition, three levels of screen brightness, a factor of demonstrated importance,[8] were specified fairly accurately.

VISUAL APPEARANCE OF THE PIP ON THE SCREEN

An earlier paper[1] in this series has described in a general way the manner in which radar system variables govern the appearance of the pip and the general features of electronic bombardment of the phosphor screen. However, no photographs were published. Photographs of small, medium and large pips are shown in Fig. 1. The photographs illustrate how the pip and the faint concentric rings, used operationally for the estimation of distance from the center of the screen, appeared to the observer. They fail, however, to portray the circular gradient of brightness during the afterglow of the rotating bright radial "sweep" line.

The long radial sweep, seen at the bottom of each photograph, rotated at a rate of 20 revolutions per minute. In actuality, 1000 separate spokes per second were generated, so that at this rate of rotation there were 8.33 adjacent fine spokes in rapid succession for each degree. Each spoke of light created on the screen briefly fluoresces to a level determined by the CRT grid bias, then fades at a rate depending on the brightness level of that initial flash. A pip appears as a slightly brighter overlay on each spoke; the portion of each spoke occupied was determined by the pulse length of the signal. The second dimension, the beam width, of the pip was equivalent to the number of spokes on which the pip rode. One degree of beam width means 8 or 9 spokes; ten degrees, 83 or 84 spokes; and so on. It is emphasized, too, that the spoke bearing the leading edge of a pip struck the eye some time before the one bearing the terminal edge. The time intervals involved are listed in Table I.

The visual angles subtended at the standard 12-inch viewing distance are shown also in Table I. These computations apply for pips presented at half-radius, the region used in this experiment. Photometric tests indicated that pip brightness did not significantly vary as either dimension of size was changed.

Apparatus

The apparatus used has also been described in greater detail in a previous report.[1] A brief summary here will suffice. Signals were fed by a special signal generator directly to the grid of the CRT. Before entering the tube the signals could be varied by an attenuator in one-half db steps. Thus a com-

[7] R. Payne-Scott, "The visibility of small echoes on radar PPI displays," Proc. Inst. Radio Eng. 36, 180 (1948).

[8] Williams, Bartlett, and King, "Visibility on cathode-ray tube screens: Screen brightness," J. Psych. 25, 455 (1948).

TABLE IV. Visibility in decibels for the brightest background.

Pulse length (microseconds)	Beam width (degrees) 1	2	3	4	5	10	30	60
0.5	14.5	15.5	20.0	20.0	22.0	21.5	25.0	29.0
1	21.5	22.5	28.0	29.0	28.5	29.5	32.0	33.5
2	25.5	28.5	32.0	31.5	32.0	33.0	36.5	38.0
3	28.0	32.0	35.0	35.0	36.5	36.5	39.0	39.5
4	26.5	29.5	32.5	33.5	33.5	36.0	38.5	40.5
10	28.5	30.0	33.5	35.0	37.0	34.5	40.5	44.5
20	28.5	31.0	33.5	33.5	34.5	36.0	38.5	41.0
30	30.0	31.0	32.5	35.0	36.0	37.5	40.0	43.0

plete range of signal intensities was possible, starting with a reference signal voltage of eight volts. The CRT used was a No. 7BP7 manufactured by RCA. The effective diameter of its viewing screen is six inches. Circuit voltages were carefully monitored at all times so that electrodes were at their optimal potentials. At the beginning of the experiment the tube was approximately 310 hours of service age. In this particular experiment no video noise was employed.

Subjects

Observers were two well-trained subjects each of whom had completed previously several hundred hours of similar observations.

Procedure

Observation was carried out in a very dimly lighted room. The position of the head was fixed at 12 inches viewing distance by the opening in a black felt hood which fitted snugly over the entire scope. Observations were made over a period of about two weeks in January and February, 1948.

The psychophysical method employed was a variation of the method of average error, in which the observer adjusted a near-threshold pip to his criterion of threshold visibility by a telephonic connection with the experimenter in the adjoining control room. Each of the observers furnished one such threshold reading for each of the 64 pip sizes at each of three levels of scope brightness. The 64 readings were randomized within the series for each brightness level. Each experimental session lasted about three hours, observers alternating in five to ten minute sessions. Both before and after each afternoon's observations, thresholds were taken on a few standard pip sizes as an additional check on the stability of the CRT circuits.

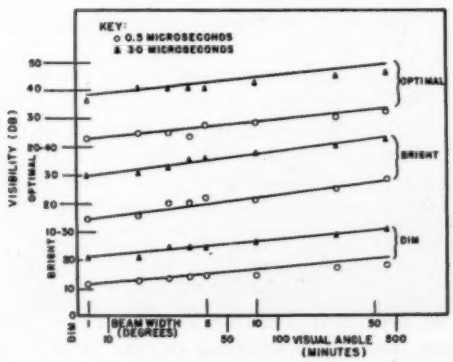

FIG. 2. Visibility as a function of the logarithm of beam width for each of two pulse lengths at each of three brightness levels. The solid lines are drawn through points determined from the equations described in the text. The indicated empirical points are taken from Tables II, III, and IV.

The three brightness levels selected for study are believed to span the normal range of operating conditions. The VRI, or Visual Reference Intensity, is a scope with a grid bias just sufficient to make the rotating sweep line barely visible to a dark adapted eye. It is referred to in the tables and figures as *Dim* and has an effective background brightness of approximately 0.0001 footlambert. This value of bias is above the "cut-off" bias listed in manufacturers' specifications. In psychological research it has proved to be a valuable reference point for background brightness. The data on pip thresholds with such a visual setting vary much less from day to day than do data with the bias set at any arbitrary constant voltage read by meters. A biasing voltage required to produce any given visual reference intensity varies considerably with tube age and excitation history. To compensate for immediate pre-excitation, the tube screen was desaturated by strong red light prior to each day's experimental session. Whenever a medium or bright scope was used the VRI was set first. The bias was then turned down to either 5½ volts (read by meter) or 10 volts. The VRI minus 5½ volts is *optimal* for visibility with this tube[8] and represents an effective background brightness of about one-tenth footlambert. The *bright* scope is VRI minus 10 volts, whose background is about 2 footlamberts.

THE MEASUREMENT OF VISIBILITY

The unit in which visibility is expressed is explained here because of its unfamiliarity to many readers. It has become quite common in radar research to express the minimum visible signal in either volts or power, for these are the variables with which the engineer must work. The visual scientist, of course, is accustomed to light units. Unfortunately there is a genuine hiatus between the two. The signal voltage is proportional to its immediate product, the current of the electron beam bombarding the phosphor of the viewing screen. Within limits, the luminous output of the phosphor screen is proportional to the beam current, but there are limits and exceptions, some of which were explained in an earlier report.[8] The dependence of luminosity on beam current (or voltage) is a function of many particular features of

tube construction and phosphor preparation, some of which are specific to individual tubes. In general, for the P7 type of phosphor used here, luminosity is approximately a linear function of voltage up to a point, after which its increment gradually diminishes. This is because of saturation effects in the phosphor; its responsiveness to electronic bombardment decreases with intensity of excitation. For this reason, a ratio of signal voltages is not always proportional to a ratio of brightnesses. As a matter of circumstance, the visibility data reported here—decibels of attenuation of a reference signal on the grid of the CRT—are roughly proportional to log contrast (or $\log \Delta I/I$) except for the dimmest scope condition used. At the present time it is next to impossible to translate accurately from electrical values to light values; although it can be done, it involves numerous assumptions as to the particular stage of phosphor decay and other variables. We have, therefore, left our data in raw form. The decibel scale is immediately meaningful to workers in radar and television and with time it may become more useful in visual circles.

RESULTS

The visibility readings obtained are presented in Tables II, III, and IV. Each entry in each table is the mean of two values, one for each observer. The table entries are decibels of attenuation of the standard signal voltage. Therefore, the higher the numerical value of the entry the greater the visibility and the lower the threshold. Visibility obviously increases with both beam width and pulse length, with the rate of increase being dependent upon the background brightness. These relationships will be demonstrated in the graphs to be presented later.

From the empirical values given in the tables, plots were made of visibility as a function of the logarithm of beam width at each brightness level for each pulse length. Since each of these functions appeared to be linear, a straight line was fitted to each set of data by the method of least squares. This procedure resulted in 24 equations, one for each of the eight pulse lengths at each of the three brightness levels. At any given brightness level the intercept constants in these equations increased regularly with increasing pulse length; this is to be expected from the values given in the tables. The slope constants, on the other hand, exhibited no such regular increase for any one background brightness. There was variation, from one pulse length to another, but with the possible exception of the dim background, the fluctuations appeared to be random and the various slope constants were not significantly different. It may be worth noting that for the dim scope, the values obtained with the four shorter pulse lengths were all lower than those with the four longer ones. The increases, however, were not systematic and are believed to represent sampling error.

Because of the insignificant differences among the slope constants for each background brightness, the mean constant was taken as the best estimate of the slope for each particular brightness. This resulted in three equations, one for each brightness, of the form,

$$\text{visibility} = a \log(\text{beam width}) + b,$$

where the slope constant a is the mean value cited above, and the intercept constant b is a function of pulse length.

Again from the data in Tables II, III, and IV, plots were drawn of visibility as a function of the logarithm of pulse length. The sets of points obtained for the eight beam widths on the dim background were again obviously linear. Unfortunately, from the standpoint of simplicity, this was not true for the two brighter backgrounds. In both of these cases the data showed distinctly curvilinear trends. In other words, visibility appears to be a simple linear function of the logarithm of pulse length, as well as of beam width for a dim background, but for brighter backgrounds a more complicated relationship is involved.

For the lowest background brightness, then, an equation of the form

$$\text{visibility} = a \log(\text{beam width}) + c \log(\text{pulse length}) + d$$

describes all the data. The least squares solution for this function gives values of 5.47, 5.56, and 12.87 for a, c, and d, respectively. Expressing the equation in more concise form, we have the following:

$$V = 5.47 \log B + 5.56 \log P + 12.87,$$

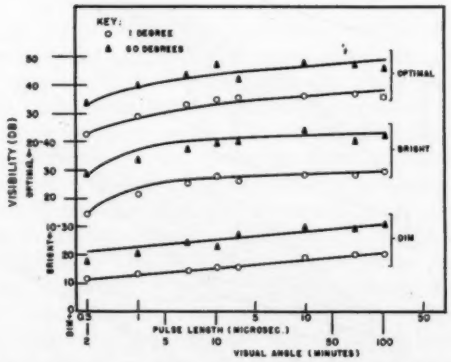

FIG. 3. Visibility as a function of the logarithm of pulse length for each of two beam widths at each of three brightness levels. The solid lines represent the empirical equations described in the text; the indicated empirical points are taken from Tables II, III, and IV.

where V = visibility in decibels attenuation of the standard signal voltage, B = beam width in degrees, P = pulse length in microseconds.

As stated earlier, the data for the two brighter backgrounds require a somewhat different function to describe them. Arbitrary constants times log log values of pulse length were needed to make the functions linear for these two conditions. The following equations were obtained as satisfactory fits.

For the optimal brightness:

$$V = 6.02 \log B + 18.67 \log \log 4P + 32.58.$$

For the brightest background:

$$V = 7.61 \log B + 8.14 \log \log (2.1)P + 27.90.$$

The symbols have the same meaning as in the equation for the dim background.

The degree of accuracy with which these equations describe the data is indicated in Figs. 2 and 3. Figure 2 shows visibility as a function of the logarithm of beam width. For the sake of simplicity only two pulse lengths, the shortest and the longest, are shown for each brightness level. The solid lines on the graph are drawn through points obtained by substitution of the appropriate values in the empirical equations listed above. The plotted points around each line represent values taken from Tables II, III, and IV. The deviations of the points from the lines are representative of the maximal discrepancies existing between experimental data and the empirical equations.

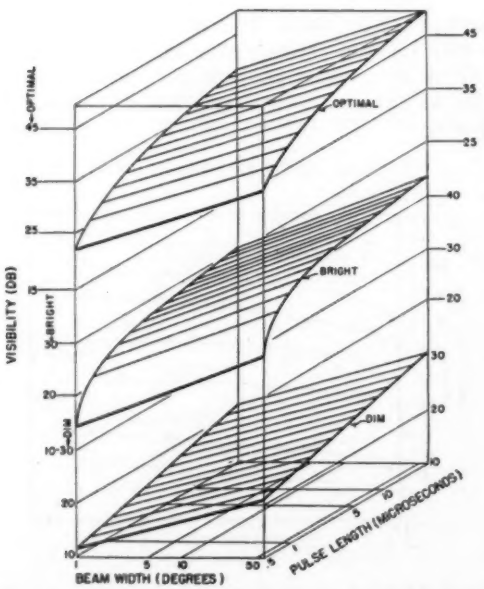

FIG. 4. Visibility as a function of beam width and pulse length for each of three brightness levels. The surfaces are drawn with values obtained from the empirical equations described in the text.

Figure 3 shows visibility as a function of the logarithm of pulse length. Again an arbitrary selection of the values to be presented was made for simplicity and clarity. Only the widest and narrowest beam widths are indicated at each brightness level. The solid lines are again drawn from the empirical equations, while the plotted points are experimental data. The curvilinear trends of the data for the two brighter backgrounds are obvious on this graph.

In both of the figures a logarithmic scale of the visual angle subtended by the various test objects is included below the beam width and pulse length scales. This allows easy transformation from one scale to the other on the graphs.

For a composite picture of the variables of the investigation, the three-dimensional graph in Fig. 4 is included. This graph illustrates the differences in the functions for the various brightness levels. For both the optimal and bright backgrounds the slope of the surface is steeper than for the dim condition, while the "bright" surface is both steeper and more curved than that for the optimal condition.

It will be recalled that no video noise was introduced for these observations. However, additional experimental data with random video noise collected at other times confirm the general results provided the scope background is uniform.

DISCUSSION

As the pip size is increased by lengthening the pulse duration or by increasing the number of successive sweeps on which the pip appears, the minimum signal voltage on the CRT grid for rendering the pip visible decreases. But in no case is the longer duration of stimulation required for increasing the size compensated by the permissible reduction in signal voltage. For example, Fig. 2 shows that a tenfold increase in the beam width for a pip on a bright screen reduces the threshold signal voltage only 7 or 8 db. A tenfold increase in beam width represents a tenfold increase in the number of sweeps, i.e., a tenfold increase in the time during which the signal is presented. And while this increase in time means that less voltage is required, still by no means is the voltage reduction of tenfold magnitude. The same point applies with the gain in visibility for increases in pulse length. The total system comprised of CRT-plus-observer falls far short of complete integration of signal energy when the signal is spread out in time, either in pulse length or beam width. The smaller the pip, the higher its signal voltage must be to render it visible; on the other hand, in terms of total electrical energy expended over the total time the signal is generated, the smallest pips are the most efficient for visibility.

With our equipment, the very smallest sizes could not be studied. Complete integration of flux for very small areas has recently been reported for rectangles on fairly bright fields.[8] For areas with either the length or width dimension subtending more than two minutes visual angle, on the other hand, integration of flux was not complete; the smaller areas were more efficient in terms of the product of area and intensity. In our experiment, all data save for one point (the smallest pulse length) apply to pips subtending more than two minutes for each dimension. It is not surprising, therefore, that our data show the smaller targets to be more efficient. In other words, we might predict from studies of the integrative properties of the visual system that, for our pips, the smallest would require the least energy.

Then too, on the basis of the persistance properties of the screen, we would expect the smallest size to be the most efficient. This is because small pips require more brightness for threshold and therefore persist for a longer time on the screen. In an earlier study[5] the duration of visibility was measured, and indeed, it was found for weak signals that there is a marked increase in duration of visibility as the signal voltage is increased. The eye can capitalize upon this resulting gain in exposure upon the screen, and thus the effect is to counteract to some extent the greater brightness requirements for a small pip. The new result of the persistence effect would be that visibility, in voltage units, would fall off less sharply with decreased pip area than we might anticipate from non-radar studies in which area alone is varied. To firmly establish this interpretation would require more data than we now have. But on present evidence, it is quite plausible that the persistence effect contributed in some measure to the greater efficiency of the smaller pips in this experiment.

In spite of the fact that the total system does not manifest complete integration when the signal is spread out in time, there *is* partial integration, i.e., the reduced voltage requirements for the larger pips are pronounced. Especially interesting is a comparison of the effects for changes in beam width with those for changes in pulse length. It will be recalled that for the dim screen the increases for the two variables showed almost the same slope constant. Indeed, as a replot of the data in Fig. 5 illustrates, all data are described satisfactorily by the equation:

$$V = 5.7 \log \text{Area} + C,$$

where V is defined as before, and C is an intercept constant. So for a dark screen, it matters little which dimension is changed; doubling one has about the same effect as doubling the other.

With brighter screens, the story is more complicated. As already stated, the relationship between visibility and log beam width is linear, but that between visibility and log pulse length is better described by a curve approaching an asymptote for long pulse lengths. We suggest that the critical difference between an increase in pulse length and an increase in beam width lies in the time factor. For example, increasing the pulse length from 10 microseconds to 20 microseconds has about the same effect on the visual angle subtended as does an increase in beam width from 5 degrees to 10 degrees, as we used the PPI in this experiment. But whereas the first change meant the pip was exposed to the eye for a total of 10 microseconds longer duration, the second change resulted in an increase of almost 42 milliseconds in exposure time. For a pip 10 degrees in beam width by 20 microseconds in pulse length the sweep transit time across the pulse length dimension is 20 microseconds, and across the beam dimension it is 83.3 milliseconds.

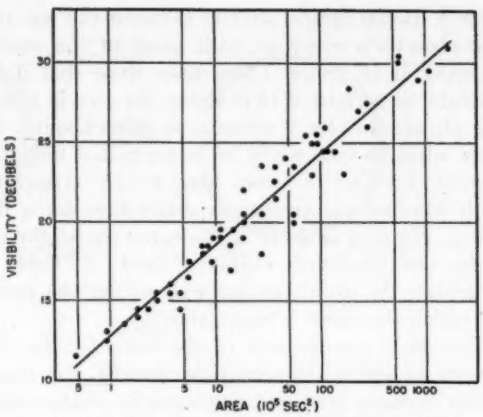

FIG. 5. Visibility as a function of area (beam width in seconds of visual angle times pulse length in seconds of visual angle) for the dim background.

One might reason that an instantaneous increase in size would permit the maximal areal integration by the eye; and as we know from vision studies, such area integration becomes progressively less effective for larger areas. An area which increases in size more slowly, on the other hand, presumably would offer less opportunity for area integration to operate. But though such integration might be less effective, the *time* during which the eye can function is increased. This last factor of statistical opportunity might be sensed over a considerable period. As a matter of fact, the point about the enhanced visibility even for increases beyond fairly long exposure times has been amply verified by Sweet and Bartlett.[9] They made tests with a PPI equipped

[9] A. L. Sweet and N. R. Bartlett, "Visibility on cathode-ray tube screens: Signals on a P7 screen exposed for different intervals," (Manuscript in preparation).

with a special optical shutter between the pip and the observer's eye, but with most of the screen showing at all times. Those tests show that if an operator is permitted to examine the area in which the pip appears for 4 seconds, he gains about 2 db over what he can see if he is permitted only 0.4 second. For our purposes here we are concerned with whether the greater exposure time for a 60-degree pip over that for a 30-degree pip might explain the improved visibility; and as Table I indicates, the durations are well within the range for which the effect is reasonable.

Obviously our account of the basis for the difference in effect between pulse length and beam width increases is an oversimplification of the visual phenomena involved. But evidently some such interpretation treating the different time factors in the two variables must be invoked.

SUMMARY

An experiment has been conducted to determine the effect upon visibility of changes in the size and shape of a radar pip. Only one type of radar presentation, the Plan Position Indicator, was used. A wide range of beam widths (which determine the angular dimension of the pip) and pulse lengths (which determine the radial dimension) was investigated at each of three test field brightnesses.

Results show that visibility, on a decibel scale, is a linear function of the logarithm of beam width at each of the three brightness levels used. For the dim background visibility is also a linear function of the logarithm of pulse length. For the brighter backgrounds, on the other hand, visibility is shown to be a curvilinear function of the logarithm of pulse length. Tables, graphs, and empirical equations are presented for summarizing the data, and relationships to more general visual problems are discussed.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

Visibility on Cathode-Ray Tube Screens: Signals on a P-7 Screen Exposed for Different Intervals*

N. R. BARTLETT** AND A. L. SWEET***
Psychological Laboratory, The Johns Hopkins University, Baltimore, Maryland
(Received February 25, 1949)

Empirical data for the minimum signal voltage for a pip on a P-7 radar screen have been described as a function of the exposure time of the pip. This has been done for two bias levels for the CRT—a dim scope and a bright scope. From the measurements it is concluded that there is a more marked gain in visibility with increased exposure time for the bright screen than for the dim one, but that on the bright screen visibility probably approaches its maximum at a shorter exposure time. Two turning rates for the sweep were used. And it may be concluded that, for any exposure time or for either brightness of screen (CRT bias) the slower turning rate yields improved signal detection.

INTRODUCTION

ONE of the important variables governing the minimum detectable contrast for a visual target presented upon a background is the duration of exposure of the target. The longer the duration, the less the minimum contrast required so it is only reasonable to suppose that a radar operator should prove better able to detect weak signal images, or "pips," on a scope screen if he is permitted a longer time to examine the area of the screen where a pip appears. But how much does he gain by increasing his time? The answer to that question is important for any theory describing the optimal conditions for observation of radar scope screens. Moreover, the answer, to be most useful, should be expressed in the engineer's language for describing the electrical signals which are fed to the input stage of the cathode-ray tube (CRT).

This paper furnishes some data for the problem as it applies to observation on a standard Plan Position Indicator (PPI) with a P-7 screen. The data show how the minimum voltage for a signal of a particular type varies with the time permitted for observation. No attempt has been made to correlate the empirical findings with data for any related studies of the minimum visible contrast. Because of its peculiar time constants the phosphor screen furnishing the light presents some difficulties in translating back and forth between signal voltages and the contrast and time units for the pip. So while our data are useful in that they provide a direct answer for the engineer on the time of

* This research was carried out under Contract N5ori-166, Task Order I, between Special Devices Center, ONR, and The Johns Hopkins University. This is Report No. 166-I-83, Project Designation No. NR-784-001, under that contract.
** Now at Department of Psychology, Hobart and William Smith Colleges, Geneva, New York.
*** Now at Psychology Section, Radio Division III, Naval Research Laboratory, Washington 20, D. C.

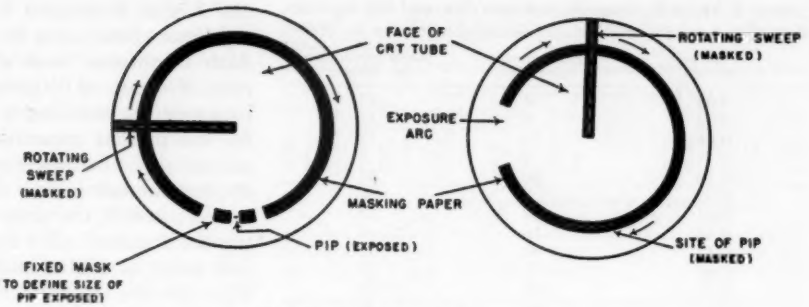

FIG. 1. Illustration of mechanical technique for varying duration of exposure of the pip (frontal view).

A. View when pip is exposed. B. View when pip is masked.

observation problem, they are notably deficient in affording any comparison with comparable studies from vision laboratories.

In our experiments the time that the pip was exposed was varied, and for each exposure time the minimum detectable signal voltage to render the pip visible was determined. This has been done over a fairly wide range of exposure times for each of two levels of screen brightness.

APPARATUS AND METHOD

General

The scope, signal generation equipment, voltage monitoring, and general method of observation have been described in earlier reports of this series[1,2] and need not be detailed here. For this experiment the reference signal voltage was defined as in a previous study[2] with the single exception in the detail that 30 db attentuation rather than 28 was used for the calibration of the standard signal. Pulse repetition frequency was 600 per second, pulse length 2 microseconds, antenna beam width approximately 20 degrees, and the range scale for the 7-inch scope was 20 miles. Two speeds of sweep rotation, 6 and 32 r.p.m., were used. No video noise was introduced.

The 20-degree pip, of which only the center 4 degrees were exposed by the apparatus described below, was presented at a position below the center of the screen and approximately $\frac{3}{4}$ of the distance out to the periphery.

Exposure Control

The apparatus for controlling the exposure time for the pip was designed to isolate the time factor while leaving all other conditions undisturbed insofar as possible. The equipment constructed for a previous study[2] was modified into the apparatus illustrated schematically in the two views presented in Fig. 1. Represented there in each view are two vertical planes. The inner (lower) one is the PPI scope face. On this screen appear (1) the rotating sweep line and (2) the pip formed by exciting a portion of the screen as the sweep passes by. Sweep and pip are shown schematically by dotted lines. On the glass of the scope face two small strips of opaque paper were attached so as to

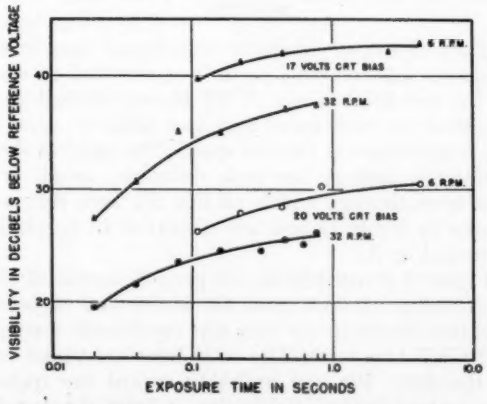

FIG. 2. Visibility, at different times of exposures, for two CRT biases and two rotational speeds.

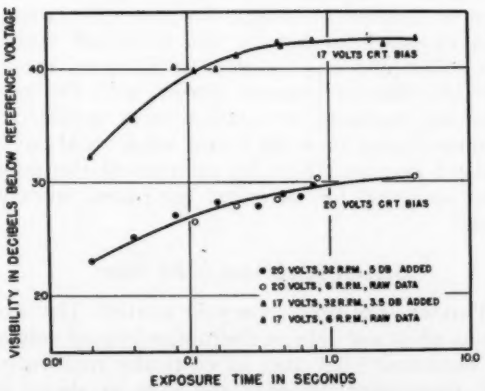

FIG. 3. Presumed form of the relationship between visibility and exposure time for 7BP7 scope operated with sweep rotating at 6 r.p.m. for two CRT bias levels.

[1] Williams, Bartlett, and King, J. Psych. 25, 455–466 (1948).
[2] A. L. Sweet and N. R. Bartlett, J. Opt. Soc. Am. 38, 329–337 (1948).

TABLE I. Times of exposure, exposure arcs, and turning rates.

Time in seconds	Exposure arc in degrees	Sweep revolutions per minute
0.02	4	32
0.04	8	32
0.08	16	32
0.11	4	6
0.16	30	32
0.22	8	6
0.31	60	32
0.44	16	6
0.47	90	32
0.62	120	32
0.78	150	32
0.83	30	6
1.67	60	6
2.50	90	6
3.33	120	6
4.16	150	6

occlude the pip save for its middle four degrees. This arrangement will be referred to hereafter as the "fixed window." Above the scope face, as a second plane, is mounted a transparent lucite disk. This disk is made to rotate at the same rate as the sweep. On the rotating disk are fastened two masks of opaque paper. One is a long thin rectangle running from the center to the periphery to mask the sweep line; the other is an annulus, $\frac{5}{16}$ inch wide, which masks the pip except when the disk is turned to expose the pip through a break in the annulus. The width of the break could be varied to yield several "exposure arcs."

The pip is masked from view, then, except once for each revolution of the lucite disk when the exposure arc passes over the 4-degree fixed window. At a constant rotational speed, the duration of the exposure is directly proportional to the length of the exposure arc. Thus, e.g., at a speed of 15 r.p.m., an arc of 45 degrees will disclose the signal image for one-half second. We capitalized upon this means for varying the exposure time.

We have explored the effect of exposure time over a considerable range. To cover the range it was necessary to employ two rotational speeds: one at 6 r.p.m. and the other at 32 r.p.m. Table I lists the times of exposure studies, with the corresponding exposure arcs and turning speeds. Exposures shorter than the lowest value could not be realized because 32 r.p.m. approaches the upper limit on speed for the radar equipment we were using.

Screen Brightness (CRT Bias)

Two levels of brightness were studied. The level has an effect not only on the minimal signal voltage for detection[1,2] but also, of particular concern for this experiment, governs the rate of decay of visibility after fluorescence.[2] One level of CRT bias chosen was 20 volts, a value which approximates the Visual Reference Intensity referred to as a significant base value in an earlier note.[1] The initial flash brightness level of the sweep line and the rate of decay of brightness afterward for such a bias together combine to offer an interesting feature for the present experiment. During the first half-second after fluorescence, visibility drops 7 or 8 db, but thereafter over the course of the next eight or ten seconds, the drop is not appreciable. During the third second after fluorescence a faint signal is just about as visible as during the previous second. Thus at the 20-volts bias, provided the pip be occluded during the first half-second of decay, increases in the exposure time over some range do not involve a simultaneous drop-off in visibility for the pip. The other level of bias, three volts below this, was chosen for study because it represents a brighter screen and near-optimal visibility for this particular tube.[1] At this lower bias there is a pronounced fall-off in visibility with decay over several seconds, and so increasing the exposure time permits a longer view, but a view of a pip fading rapidly into the background.

Determination of Thresholds

For each datum the experimenter began by presenting a weak signal, and then decreased the attenuation in $\frac{1}{2}$ db steps until finally *S* phoned that he detected the pip on the screen. The method was crude, but the data showed a satisfying reproducibility. Troublesome persistence effects from phosphorescence would have resulted, particularly with the higher bias, had we used a descending order from supra-threshold to invisible. So the method adopted was a compromise of the psychophysical method of limits with the equipment realities.

RESULTS

Each of the five subjects contributed data in six separate sessions. One series was devoted to each of the two brightnesses (CRT biases) studied with the slow rotation speed and two series to each of the brightnesses at the fast speed. The mean of their threshold readings for each duration, speed, and bias level through which all five *S*'s were run, are shown in Table II and are displayed in graphical form in Fig. 2.

Figure 2 demonstrates the general trends of the relationship between mean thresholds and exposure duration shown in the four test conditions of speed and CRT bias level. The solid lines are visual fits to the data. We can probably regard the trends depicted as fairly reliable, for at least the trends shown by these means for each of the four conditions is fairly apparent in the data for each of the five subjects; there is substantial agreement in the data for the five.

From Table II and Fig. 2 we can come to a few conclusions. First, the point of the superior visibility at the 17 volts bias over that at the 20 volts bias reported in previous articles[1,2] is confirmed once again. Second, the gain in visibility with increased exposure time is evidently more pronounced at the higher brightness level (lower CRT bias). In proceeding from exposures of 0.02 seconds to one second the gain (with fast sweep rotation) is approximately 10 decibels for the brighter scope, whereas for the same increase in time for the dim scope the increase is only some 6 decibels. This finding, moreover, means greater improvement despite the fact that on the brighter scope there is a more rapid fading of visibility after fluorescence.[2] Finally, an increased visibility for a particular bias level at the slower sweep speed seems fairly established. As a guess we attribute the greater attenuation possible with the slower speed to a slightly enhanced brightness of the scope screen for the slower speed. In all likelihood the threshold for signal detection depends upon sweep speed, and, moreover, the bias level which is optimal for detection of weak signals is also dependent in some degree on sweep rotational rate.[1]

The data for the fast rotation shown in Fig. 2 have been adjusted by an empirical correction so as to illustrate, in Fig. 3, the probable relationship between exposure time and mean thresholds for a scope operated with its sweep turning at 6 r.p.m. We have assumed that the 32 r.p.m. data at the 20 volts bias fall along the same curve as do the 6 r.p.m. data, but at a level 3.5 db lower. Similarly we assumed that the 32 r.p.m. data at the 17 volts bias are 5 db lower. Accordingly, we added 3.5 db to all the 32 r.p.m. data taken at 20 volts, and 5 db to all the 32 r.p.m. data taken at 17 volts. The results after these empirical corrections are illustrated in Fig. 3.

TABLE II. Mean threshold readings for five subjects, expressed as number of decibels of attenuation for a reference signal voltage.

	Durations of exposure—sweep speed 32 r.p.m.*							
CRT Bias	0.02	0.04	0.08	0.16	0.31	0.47	0.62	0.78
20 Volts	19.5	21.5	23.5	24.7	24.3	25.4	25.1	26.2
17 Volts	27.2	30.6	35.2	35.0		37.0		37.4

	Durations of exposure—sweep speed 6 r.p.m.**					
	0.11	0.22	0.44	0.83	2.50	4.16
20 Volts	26.3	27.9	28.4	30.2		30.5
17 Volts	39.8	41.2	42.0		42.2	42.8

* Each *S* measured once on each of two days.
** Each *S* measured once only.

If our compromises with the data are warranted, then the curves, fitted by visual inspection, are approximations for the relationship between exposure time and the minimal detectable signal voltage for the P-7 screen with a 6 r.p.m. sweep. And from these curves we may make some reasonable guesses about the optimal exposure time. For the bright screen (17 volts CRT bias), the gain in visibility with increased exposure time is first rapid, but the curve approaches a maximum fairly soon. An exposure of one-fifth second is probably within 1 or 2 db of the level of performance an operator might manifest with an exposure of several seconds, and even with only one-tenth second he is probably within 3 db of his level for a very long exposure. On the other hand, for the dim scope (20 volts CRT bias), the gain in visibility with increased exposure time is much more gradual. Decreasing the exposure time below one-fifth second does not handicap the operator nearly so much as a similar decrease with the brighter screen, and on the other hand, there is some suggestion that there may be an appreciable advantage in very long exposures of two or three seconds.

The Spectra of Argon, Krypton, and Xenon between 1.2 and 2.2 Microns*

W. R. SITTNER** AND E. R. PECK
Department of Physics, Northwestern University, Evanston, Illinois
(Received February 28, 1949)

The spectrum of argon, krypton, and xenon excited by a repetitive condenser discharge method has been studied in the near infra-red region. The use of this method of excitation and a large infra-red spectrograph using a lead sulfide detector reveals about 170 new lines in these gases. About three-fourths of the lines are classified by known terms of the normal atom.

Some anomalies of spectra excited by the repetitive condenser discharge method are mentioned.

I. INTRODUCTION

THE near infra-red spectra of argon, krypton, and xenon have been recorded by means of the grating spectrometer with a lead sulfide detector described by Nelson.[1] Wave-length was determined from the spectrum charts by means of a quadratic formula relating wave-length and a series of marks impressed on the chart simultaneously with the spectrum. The constants of the formula were determined for each chart individually by means of overlapping order or by computation of the wave-length from the term values if the transition for a first-order line was already known. Vacuum wave numbers were calculated from observed wave-length by extrapolation of the formulas of Meggers and Peters[2] for refractive index as a function of wave-length. These formulas also include corrections for the effect of temperature and water vapor. The transition has been assigned to a sufficient number of lines of argon and krypton in this region[3–5] to establish, in some cases, a correction curve which was then applied to the new lines. The maximum correction was 0.4 cm^{-1} for argon and 0.9 cm^{-1} for krypton.

II. SOURCE AND METHOD OF EXCITATION

The method of excitation is by a repetitive condenser discharge through the gas. A condenser is charged to a high potential through a resistor by a power supply. The source geometry and pressure of gas filling are such that breakdown does not occur at the charging potential. High potential pulses are applied to an exterior triggering lead around the source. These pulses cause breakdown of the gas, thus releasing the energy stored in the condenser within the gas in a few microseconds and producing a brilliant flash of light. Triggering pulses are applied at approximately 20 per second, while the charging time of the condenser is less than a millisecond, and thus the potential at the time of discharge is controlled by the power supply. Typical values of potential and capacitance are 1 microfarad charged to 1.5 kilovolts.

The electrical frequency spectrum of the undispersed flashes of light was determined by means of an harmonic analyzer. The frequency response characteristics of lead sulfide detectors of the type used in this spectrometer have been reported by Cashman.[6] From the above information a broad band amplifier suitable for recording "flash" spectra was constructed. The range of the amplifier is from 300–15,000 cycles $sec.^{-1}$

The sources are constructed with electrodes attached to the side of a long tube used for end-on viewing. The tube is 18 cm in length and 7 mm inside diameter. The diameter is chosen so that the reduced image on the detector is equal in length to the area of highest sensitivity. The sources are filled to a pressure of about 50 mm with the spectroscopically pure grade of gas supplied by the Linde Air Products Company.

III. RESULTS

The results are given in Tables I, II, and III for argon, krypton, and xenon, respectively. The columns list the mean value of vacuum wave number observed, ν°_0; the transition believed to give rise to the line; the calculated wave number for this transition based on term values given by Humphreys and Meggers,[7] ν_c°; the wave number difference in measured and calculated values, $\Delta\nu^\circ$; the approximate intensity of the line, I; and the references to previous observers, if any. The numbers in this column are those of the numbered references.

TABLE I, TABLE II, AND TABLE III

The classifications given above are in accordance with the usual selection rules except that approxi-

* This work was supported by the Bureau of Ships, U. S. Navy.
** Now at Bell Telephone Laboratories, Murray Hill, New Jersey.
[1] R. C. Nelson, J. Opt. Soc. Am. 39, 68 (1949).
[2] W. F. Meggers and C. G. Peters, Scientific Papers of the Bureau of Standards, No. 327 (October 31, 1918).
[3] R. M. Woods and B. J. Spence, Phys. Rev. 45, 669 (1934).
[4] C. J. Humphreys and E. K. Plyler, Bur. Stand. J. Research 38, 499 (1947).
[5] W. F. Meggers, Bur. Stand. J. Research 14, 487 (1935).
[6] R. J. Cashman, J. Soc. Mot. Pict. Eng. 49, 342 (1947).
[7] C. J. Humphreys and W. F. Meggers, Bur. Stand. J. Research 10, 139 (1933); 10, 427 (1933).

TABLE I. Infra-red spectra of argon.

$\nu°_{obs.}$	Transition	$\nu°_{calc.}$	$\Delta\nu°$	I	
cm^{-1}		cm^{-1}			
8718.00	$2p_7-3s_1''$	8717.85	0.15	30	5
02.20	$1s_2-2p_{10}$	8702.20	0	40	4
8633.05	$3d_4-4W$	8632.91	+0.14	10	5
00.81				4	
8567.56	$2p_6-3s_1''$	8567.56	0	50	4
53.84	$2p_7-3s_1'''$	8553.74	0.10	10	5
30.29	$2p_8-3d_2$	8530.29	0	20	4
20.33	$3d_4-4X$	8520.59	−0.26	20	5
8412.43	$3d_3-4Y$	8411.85	0.58	20	5
03.22	$2p_9-3s_1'''$	8403.45	0.23	10	5
8370.63	$3d_5-4X$	8370.49	0.14	50	5
12.85	$2p_8-3s_1'$	8312.57	0.28	20	5
8254.06	$2p_9-3d_1'$	8254.06	0	60	5
35.22	$2p_4-3s_2$	8235.15	0.07	40	5
27.33	$3d_1''-4W$	18227.32	10.01	30	5
8099.40	$2p_8-3d_1'$	8099.49	−0.09	50	5
90.87	$3d_3-4Y$	8090.87	0	50	5
73.76				5	
60.44	$2p_7-3d_2$	8060.36	0.08	60	3
49.95	$3d_3-4X$	8049.51	0.44	30	5
37.02	$2p_{10}-3d_3$	8037.00	0.02	60	3
25.83	$2p_8-2s_4$	8025.82	0.01	50	5
05.68	$2p_9-2s_5$	8005.65	0.03	60	4
7962.87	$2p_9-3d_1''$	7963.28	−0.41	10	5
36.60	$3d_1'-4W$	7936.54	−0.06	7	
20.09	$2p_5-2s_2$	7920.76	−0.67	7	5
09.21	$2p_6-3d_2$	7910.07	−0.86	2	
7870.46	$2p_2-3s_1'$	7870.44	0.02	50	3
50.91	$2p_3-2s_5$	7851.08	−0.17	33	5
43.09	$2p_4-2s_2$	7843.34	−0.25	30	5
08.75	$2p_3-3d_1''$	7808.71	0.04	60	3
7729.26	$2p_4-2s_3$	7729.85	−0.59	40	5
15.92	$2p_{10}-3d_5$	7716.02	−0.10	60	3
7684.99	$2p_3-2s_2$	7685.35	−0.36	50	3
73.28				2	
7565.11	$2p_{10}-3d_4$	7565.92	−0.81	50	
57.49	$2p_9-3d_4$	7557.69	−0.20	50	3
55.78	$2p_7-2s_4$	7555.89	−0.11	39	
7532.27				60	3
15.34				5	
09.41	$2p_4-3s_1'''$	7509.29	0.12	60	3
7499.67	$2p_4-7s_1''$	7498.94	0.73	30	
79.06	$2p_5-3d_1'$	7479.27	−0.21	60	3
56.92				50	
05.45	$2p_5-2s_4$	7405.60	−0.15	20	
03.21	$2p_3-3d_4$	7403.12	0.09	60	3

TABLE I.—*Continued.*

$\nu°_{obs.}$	Transition	$\nu°_{calc.}$	$\Delta\nu°$	I	
cm^{-1}		cm^{-1}			
7380.74				4	
64.47	$2p_2-2s_3$	7365.14	−0.67	20	
51.02	$2p_3-3s_1'''$	7351.30	−0.28	40	
38.74	$2p_7-3d_1''$	7338.78	−0.04	50	3
08.83	$2p_2-3s_1''$	7308.69	0.14	50	3
7287.54	$2p_9-3d_4'$	7287.60	−0.06	60	3
29.85	$2p_6-2s_5$	7230.86	−1.01	30	
08.92	$3d_4-4Y$	7209.46	−0.54	2	
7187.34	$2p_6-3d_1''$	7188.49	−1.15	50	
44.70	$2p_2-3s_1'''$	7144.58	0.12	40	
7093.33	$2p_5-3d_2$	7093.33	0	50	3
15.82	$2p_4-3d_2$	7015.91	−0.09	20	
12.16				40	
6849.14	$3s_1''-4Z$	6849.39	−0.25	30	
31.49	$4d_5-6Z$	6831.81	−0.32	40	
23.98				20	
03.73	$3d_1''-4Y$	6803.87	−0.14	5	
6782.65	$2p_6-3d_4$	6782.90	−0.25	5	
60.89	$2s_5-4Y$	6761.50	−0.61	1	
21.67				1	
6644.31	$2p_1-3s_1'$	6644.23	0.08	40	
6588.82	$2p_5-2s_4$	6588.86	−0.04	20	
47.05				4	
33.62				20	
21.73	$2p_3-3d_2$	6521.71	0.02	6	
13.24	$3d_1'-4Y$	6513.09	0.15	10	
11.51	$2p_4-2s_4$	6511.44	0.07	10	
6490.84				8	
6288.42	$3s_1'-4Z$	6287.64	0.78	40	
52.32	$2p_1-2s_2$	6252.42	−0.10	30	
6178.12	$2p_3-2s_5$	6178.71	−0.59	2	
6092.12				7	
82.33	$3d_2-4Y$	6082.29	0.04	10	
	$2p_1-3s_1''$	6082.48	−0.15		
52.55	$2p_7-3d_3$	6051.78	0.77	10	
41.06	$3d_2-4X$	6040.93	0.13	6	
5971.86	$2p_2-2s_5$	5971.99	−0.13	6	
01.59	$2p_6-3d_2$	5901.49	0.10	50	3
5808.35				2	
5580.55	$2p_7-3d_5$	5580.70	−0.15	10	
	$2p_6-3d_5$	5580.51	0.04	10	

mately one-eighth of the transitions indicate a change in the state of the ion ($^2P_{\frac{1}{2}} \leftrightarrow {}^2P^{\frac{3}{2}}$). There are also two cases in Table II and one case in Table I of combination between f and s terms. It is usually considered in these cases that an f electron jumps to a p state and a p electron jumps to an s state.

It is to be noted that a portion of the new lines observed are not classified by transitions between known terms of the normal atom. The absence of observations of any impurities in the spectra of these sources (except for the well-developed Xe spectrum from the Kr sources which is apparently due to a small amount of Xe in the Kr supplied by the manufacturer), in conjunction with the fact that there are no common unclassified lines of the three gases, indicates that the unknown lines are probably characteristic of the gas itself. The unclassified lines are not explained, apparently, by terms of the singly charged ion.

One of the characteristics of the flash spectrum of the inert gases as distinguished from the spectrum of an arc, or hollow cathode, source is that many transitions arising from d or f levels are greatly enhanced in the flash spectrum. Therefore, a systematic attempt was made to determine if the unclassified lines are due to unknown f terms. The conclusion is that it is not possible to make a statement regarding the existence of new terms in these gases on the basis of data of the present precision.

A distribution curve of the values of $\Delta\nu°$ for each of the gases indicates that the most probable difference in observed and calculated wave number is approximately 0.1 cm^{-1}. A number of the values are much greater than this but these are generally weak or broad lines.

TABLE II. Infra-red spectra of krypton.

ν^0obs.	Transition	ν^0calc.	$\Delta\nu^0$	I	
8277.96	$3d_3-4Z$	8277.62	0.34	50	5
50.09	$3d_4'-4W$	8250.09	0	40	4
45.82	$1s_2-2p_8$	8246.15	−0.33	5	5
8191.13	$3d_4'-4U$	8191.43	−0.30	70	
13.83				30	
7821.12	$3d_4-4W$	7821.17	−0.05	40	
7794.70	$3d_4-4Y$	7794.99	−0.29	10	
72.41	$1s_3-2p_7$	7772.77	−0.36	10	
62.56	$3d_4-4U$	7762.51	0.05	50	4
41.05				3	
7698.32	$2p_{10}-3d_1''$	7699.01	−0.69	10	
76.47				50	
7586.66	$2p_8-2s_4$	7586.73	−0.07	50	4
78.37				46	
66.35				4	
49.90	$2p_4-3s_1'$	7550.43	−0.53	50	
7494.87				30	
7154.11	$3d_1''-4Y$	7153.82	0.29	10	
6968.53	$3d_1'-4W$	6968.17	0.36	50	
6941.89	$2p_2-2s_2$	6942.14	−0.25	30	4
6929.97	$2p_7-2s_4$	6929.72	0.25	50	
6909.74	$3d_1'-4U$	6909.51	0.23	40	
6886.32	$3d_1'-4Z$	6885.97	0.35	5	
6796.29				4	
6772.03	$2p_8-3d_1'$	6771.97	0.06	60	4
6770.39	$2p_6-2s_4$	6770.77	−0.38	60	
6681.56	$2p_7-3d_2$	6681.29	0.27	50	
6661.74	$2p_7-2s_5$	6662.55	−0.81	30	
6587.65				20	
6573.35	$2p_9-3d_1''$	6573.12	0.23	20	
6560.05	$2p_8-3d_1''$	6560.14	−0.09	60	
6548.75				3	
6529.86				3	
6522.53	$2p_8-3d_2$	6522.34	0.19	5	
6519.19	$2p_{10}-3d_3$	6519.19	0	60	4
6503.09	$2p_6-2s_3$	6503.60	−0.51	50	
6460.84	$2s_2-2p_8$	6460.67	0.17	30	
6394.20	$2s_5-4Y$	6394.40	−0.20	3	

TABLE II.—*Continued.*

ν^0obs.	Transition	ν^0calc.	$\Delta\nu^0$	I	
6375.59	$3d_2-4Y$	6375.66	−0.07	40	
6319.40	$3d_2-4X$	6318.84	0.56	20	4
	$3d_2-4Z$	6319.64	−0.24		
6290.53	$2p_1-3s_1'$	6291.27	−0.74	7	
6227.31	$3d_3-3p_7$	6228.44	−1.13	6	
6127.23	$2s_4-4Y$	6127.23	0	2	
6032.59	$2p_1-2s_2$	6032.24	0.35	2	
5982.33				50	
5979.64				2	
5976.46	$1s_2-2p_{10}$	5976.89	−0.43	20	
5965.93				1	
5956.05	$2p_8-3d_1'$	5956.01	0.04	60	
5931.86	$2p_9-3d_4$	5931.95	−0.09	40	4
5928.73				1	
5926.33				1	
5918.95	$2p_8-3d_4$	5918.97	−0.02	60	
5916.75	$2p_{10}-3d_5$	5916.53	0.22	50	
5908.98				1	
5903.06	$2p_7-3d_1''$	5903.13	−0.07	60	
5856.69				1	
5846.91	$2p_4-3s_1''$	5846.77	0.14	40	
5770.12				1	
5756.36	$2p_2-3s_1'''$	5756.26	0.10	50	4
5744.15	$2p_6-3d_1''$	5744.18	−0.03	3	
5674.22	$3d_3-3p_6$	5674.92	−0.70	1	
5602.97	$2p_{10}-3d_6$	5602.99	−0.02	40	4
5553.34	$2p_6-3d_2$	5552.82	0.52	50	
5523.50	$2p_3-3s_1''$	5523.54	−0.04	2	
5502.93	$2p_9-3d_4'$	5503.03	−0.10	60	4
5497.57	$2p_2-3s_1''$	5497.52	0.05	4	
5380.45	$2p_6-3d_3$	5380.32	0.13	3	
5346.76	$2p_3-3s_1'''$	5347.25	−0.49	10	
5321.61	$2p_2-3s_1'''$	5321.23	0.38	7	4
	$1s_2-2p_{10}$	5321.80	−0.19		
5317.88	$3d_4'-3p_9$	5318.22	−0.34	3	
5317.17				3	
4896.06	$3p_8-5d_2$	4896.17	−0.11	1	
4566.99				8	

A small section of a record typical of those made in this experiment is shown in Fig. 1. The scanning rate and slit width are typical of those used in this experiment. A number of the lines of this section are not present in the spectrum excited in a hollow cathode source. The width of the lines of this section, measured at half amplitude, ranges from 0.1A to 2.5A.

IV. ANOMALIES OF THE FLASH SPECTRUM

Since there have been few studies of spectra excited in the manner here described, a few of the unusual characteristics of such spectra are mentioned.

The spectrum of a flash source appears to depend more on the energy released in the gas per flash per unit length of the discharge than on the other conditions of operation. This quantity will

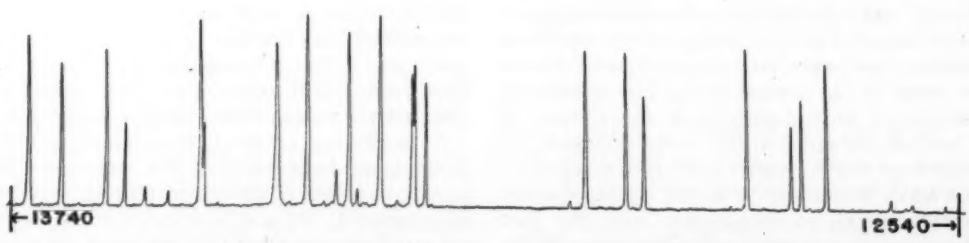

FIG. 1. Argon spectrum from flash source operated at 1.4 kv and 0.35 μf. Scanning rate, 50A/min.; slit width 2A.

be called energy "density" for brevity. At low energy density (0.02 joule/flash cm), the spectrum resembles the usual arc spectrum. As the energy density is increased the spectrum changes continuously until at high energy density (0.40 joule/flash cm) the near infra-red spectrum is characterized by:

(1) An intense continuum, which accounts for practically all of the visible radiation, extending into the infra-red but decreasing in intensity with increasing wave-length. As the energy density is increased the continuum extends to longer wave-length. This continuum, which occurs during and after the heavy current surge following breakdown, likely results from recombination of electrons and ions.

(2) A number of lines which are not excited to a comparable extent in any source known to the writers. Many of these lines arise from *d* and *f* terms of the atom. It has been observed previously[8] that higher series members of several atoms are relatively more intense when observed with recombination than when excited by other methods.

(3) All of the spectral lines being greatly broadened. At the energy density mentioned above typical widths at half amplitude are 6–12A. Generally speaking, those lines arising from *f* terms are broader than those from *d* terms which, in turn, are broader than those from *p* terms. The broadening is probably due to Stark effect in the strong inhomogeneous electric field in the region of high ion density. It is known[9] that first-order Stark effect in complex spectra is shown only by non-penetrating orbits, and this is in accordance with our observation of preferential broadening.

In order to locate as many new lines as possible, but also to make precise measurements on the stronger lines, a range of energy densities has been used in these experiments.

TABLE III. Infra-red spectra of xenon.

ν°obs.	Transition	ν°calc.	$\Delta\nu^\circ$	I	
9683.54				20	
9175.74	$3d_3-2p_3$	9175.72	0.02	60	5
8839.48	$3d_2-2p_3$	8839.60	0.12	2	5
8758.06	$3d_3-3p_7$	8757.92	0.14	6	5
8699.94	$3d_3-3p_8$	8699.86	0.08	2	5
8607.80	$3d_6-2p_4$	8607.849	−0.05	30	5
8514.02	$3d_1'-4W$	8513.84	0.18	70	5
8476.84	$3d_1'-4U$	8476.89	−0.05	50	5
8431.47	$3d_1'-4Z$	8431.34	0.13	50	5
8418.53	$3d_1'-4Y$	8419.25	−0.72	5	5
8392.50	$3d_5-2p_4$	8392.49	0.01	8	5
8363.26	$3d_3-3p_6$	8363.74	−0.48	6	5
8272.59	$3d_4'-3p_8$	8272.57	0.02	15	5
8192.10	$3d_4-2p_3$	8191.90	0.20	5	
8170.88	$2p_{10}-2s_4$	8170.88	0	50	5
8155.85	$3d_6-3p_{10}$	8155.85	0	8	5
8145.82	$3d_3-3p_8$	8146.45	−0.63	2	
8028.32	$3d_2-3p_9$	8028.92	−0.60	2	
7940.70	$3d_5-3p_{10}$	7940.49	0.21	20	
7919.30	$2p_{10}-2s_5$	7919.66	0.36	100	5
7320.19	$2p_9-2s_4$	7320.23	−0.04	70	
7068.95	$2p_9-2s_5$	7069.01	−0.06	50	
7020.18	$3d_2-4V$	7020.11	0.07	60	
6959.46	$3d_2-4Y$	6959.50	−0.04	50	
6949.89	$3d_2-4X$	6949.83	0.06	3	
6819.02	$3d_1''-3p_7$	6818.96	0.06	5	
6785.70	$2p_8-2s_5$	6785.75	−0.05	200	
6484.16	$3p_7-2s_4$	6483.99	0.17	70	
6227.44	$2p_8-2s_4$	6227.56	−0.12	20	
5976.05	$2p_8-2s_5$	5976.34	−0.29	20	
5770.24	$2p_9-3d_2$	5770.17	0.07	5	
5320.79	$2p_5-2s_4$	5321.06	−0.27	3	
5320.06				3	
5268.57				3	
5267.49				3	
5135.58				4	
4936.37				2	
4933.96	$2p_7-3d_2$	4933.93	.03	6	
4749.56				2	

V. ACKNOWLEDGMENTS

We wish to express our thanks to Mr. W. R. Wilson for the design of the broad band amplifier for recording flash spectra, to Mr. C. F. Hendee for constructing and filling the sources, and to Mr. H. B. Billington for assistance with the computations. Especial thanks are also due Dr. W. S. Huxford, director of the project under which the work was done.

[8] See references in Ruark and Urey, *Atoms, Molecules and Quanta* (McGraw-Hill Book Company, Inc., New York, 1930), p. 470.

[9] Ruark and Urey, *Atoms, Molecules and Quanta* (McGraw-Hill Book Company, Inc., New York, 1930), p. 343.

The Influence of Extraneous Elements on Line Intensity. IV. Extraneous Element Effects in the Direct-Current Arc*

WALLACE R. BRODE** AND DONALD L. TIMMA***
Chemistry Department, Ohio State University, Columbus, Ohio
(Received March 2, 1949)

The effect of varying amounts of extraneous element upon the line intensity of different elements has been investigated. Slavin's total energy method was used to make possible a more direct correlation than has been achieved using the internal standard method. The results of this study indicate that the effect of the extraneous element varies with the amount of extraneous element, but not in a simple linear fashion. The effect of an extraneous element on a group of lines of an element ending at the same energy level or beginning at the same energy level, is the same.

The experimental results show that the elements can be arranged in a series bases on their inter-element effects in the d.c. arc. An arrangement of the elements in a series based on their boiling points and excitation potentials is in agreement with that found experimentally.

THE frequencies of the lines in the spectrum of a sample indicate its composition qualitatively while the relative intensities of the lines have been shown to be a function of the quantitative composition.

The frequencies of the lines (or their wavelengths) are a fundamental property of the atom involved and cannot be influenced to any great extent by the physical surroundings of the atom. However, the spectral line intensity of an element has been shown to be influenced by the other materials which make up the sample.[1] This effect is known as the "extraneous element effect" and numerous reports of its existence have been made and some studies have been undertaken to evaluate it. Because of the dependence of the observed intensity upon factors other than concentration alone a more complete and thorough understanding of these effects becomes imperative for the further refining and improving of quantitative spectrochemical analysis.

This study was undertaken with the object of evaluating quantitatively the effect of different cations upon the line intensities of other elements. By using samples of very simple composition it was possible to correlate the observed variation in line intensity of the test element with changes in the amount of extraneous element present in the sample. Study of an extended series of such investigations indicated that the elements could be arranged in a spectrochemical series such that any element would cause a repression of the line intensity of other elements below it in the series and an enhancement of the intensity of the lines above it. A similar arrangement of the elements based upon their physical properties was found to agree quite well with the experimentally determined order.

THEORY

It is generally accepted that in the direct-current arc the excitation actually takes place in the discharge column. Therefore any factors which might affect the volatilization of the atoms into the arc will also affect its intensity, thus indicating that the boiling point of the sample in comparison to the temperature of the arc becomes an important factor. Once the atoms have been vaporized into the arc column enough energy must be supplied to the atom to cause excitation. The fraction of the total number of atoms of a given kind which will receive that amount of energy will depend upon many factors such as temperature, electrode shape, and air currents across the discharge. Consequently for a given set of electrode conditions there exists what might be defined as an effective exciting power and a comparison of this effective exciting power with the excitation potential of the atomic species under investigation would indicate the probability of an atom being excited when volatilized into the arc. Therefore the intensity, I, of the lines of an element should be proportional to two factors, one representing the ability of the arc to vaporize the sample and second the ability of the arc to excite the atoms. This is expressed by the following equation:

$$I_i \alpha (k - Bp_i)(h - E_i), \qquad (1)$$

* An abstract of a dissertation presented in partial fulfillment of the requirements for the degree Doctor of Philosophy at Ohio State University by Donald L. Timma.

** Now Associate Director, National Bureau of Standards.

*** Now at Mound Laboratories, Monsanto Chemical Company, Miamisburg, Ohio.

[1] W. R. Brode and E. S. Hodge, J. Opt. Soc. Am. **31**, 58 (1941); W. R. Brode and J. L. Appleton. *Proceedings of The Seventh Summer Conference on Spectroscopy and Its Applications*, (John Wiley and Sons, Inc., New York, 1940) p. 36; W. R. Brode and R. W. Silverthorn, *Proceedings of The Sixth Summer Conference on Spectroscopy and Its Application* (John Wiley and Sons, Inc., New York, 1939) p. 60; R. L. Mitchell and I. M. Robertson, J. Soc. Chem. Ind. **55**, 269 (1936); Victor R. Ells, J. Opt. Soc. Am. **31**, 534 (1941); Sawyer, Waggoner, and Erickson, *Proceedings of the Seventh Summer Conference on Spectroscopy and Its Applications*, (John Wiley and Sons, Inc., 1940) p. 47; G. O. Langstroth and D. Andrychuk, Can. J. Research, **26A**, 39 (1948); W. Zehden, J. Soc. Chem. Ind. **59**, 236 (1940).

where I_i=intensity of the lines of element i, Bp_i=boiling point of element i (degrees centigrade), E_i=excitation potential of element i (electron volts), h=a constant of the source being a measure of its capacity to excite atoms, and k=a constant representing the temperature of the arc. To evaluate the constants h and k it is necessary to have two elements which when present in a mixture are excited at the same time by the arc and another pair which are excited simultaneously but at a different time than the first pair. Rusanov[2] in studying the properties of the arc discharge has found such pairs to be Sb−Bi and Fe−Co. Using the boiling points and excitation potentials for these elements the constants h and k have been evaluated so that

$$I_i \alpha 6720 - 1600E_i - 4.2Bp_i + E_iBp_i. \quad (2)$$

By substituting the values of the boiling points and excitation potentials into formula (2) the relative intensities of the lines of the different elements could be determined for a certain time during the discharge period. In this manner the elements could be arranged in a spectrochemical series based on the properties of the elements.

EXPERIMENTAL

In order to simplify the composition of the sample as far as possible the total energy method of Slavin[3] was adopted. This allowed the samples to be composed of only two parts—the test element present in constant amount and the extraneous element present in varying amounts. In this way it was possible to study the effect of a single element upon another element. This allows a more specific assignment of the effects than when the observed quantity is the change of an intensity ratio. Water solutions of the samples as chlorides were placed on the carbon electrodes by the use of a micropipette and the solutions evaporated to dryness. In each set one of the electrodes contained only the test element by itself while the other eight electrodes contained in addition to the test element the extraneous element in amounts ranging from 500 to 3.9 micrograms. Five such sets were prepared for each plate of exposures so that five separate determinations were made for each concentration of extraneous element for each of the test elements.

A large Hilger Littrow type spectrograph which has a dispersion of 2.95A/mm at 3100A was used for this work. A direct-current arc operated at 9.0±0.5 amperes from a motor-generator set delivering 125 volts was used for excitation. An exposure using an iron arc and a step filter was made on each plate so that a plate calibration curve could be made. Eastman spectrum analysis No. 1 plates were used throughout this study. They were developed in Eastman D-19 for 5 minutes at 20°C, placed in an acetic acid short-stop for 10 seconds, and cleared in Eastman F-5 acid-hardening fixer after which they were washed and dried.

An Applied Research Laboratories comparator-densitometer was used to measure the density of the various steps of the iron exposure for constructing the plate calibration curve, and to measure the density of the lines of the test element in each of the exposures.

It was desirable to compare the effect of the extraneous element on different lines of the same element and on lines of different elements even though the lines of a single element or the lines of different elements have different inherent intensities. As a common basis for such a comparison the ratio, f, was chosen. Where f was defined as the ratio of the intensity of a line of the test element in the presence of the extraneous element to the intensity of the same line when the sample contained no extraneous element. The magnitude of the deviation of this ratio from unity is then a measure of the effect of the extraneous element. Typical curves showing the relationship between f and the extraneous element concentration for different combinations are shown in the Fig. 1.

The test elements chosen for study were Cd, Bi, Zn, Pb, and Sn. They were selected because their salts could be obtained readily in a pure form, they covered a wide range of excitation potentials, and because the electrodes used were free of impurities of any of them. The latter is particularly important when using the total-energy method because large errors are introduced by small amounts of impurities. The inter-effects of these elements were studied along with the effects of Mn, Ca, Na, Fe, Al, Cu, and V upon each of them.

APPLICATION

If in the absence of an extraneous element the intensity of a line of the unknown element is I_1

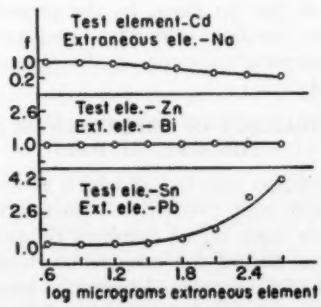

FIG. 1.

[2] G. S. Smith, Spectrochimica Acta 3, 235 (1948).
[3] M. Slavin, Ind. Eng. Chem. An. Ed. 10, 407 (1938).

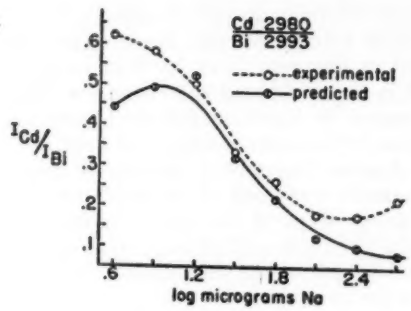

FIG. 2.

and the intensity of the internal standard line is I_2 then the intensity ratio, R, is given by $R=I_1/I_2$. Now if an extraneous element is added the intensity of the unknown element line is changed and becomes I_1f_1 and that of the internal standard line becomes I_2f_2 and the new intensity ratio observed is $R'=I_1f_1/I_2f_2=I_1/I_2\times f_1/f_2$. The values of f_1 and f_2 are obtained from the preceding experimental data and the change in R' can be predicted for variation in the amount of extraneous element. The result of such a study showing the effect of sodium (extraneous element) on Cd/Bi is shown in Fig. 2. Further studies showed that the effect of calcium on Cd/Sn and lead on Cd/Sn could be predicted in a like manner from the experimentally determined factors.

An investigation of the energy levels involved for the various lines studied showed that in general a group of lines of an element which results from transitions beginning at the same energy level or which end at the same energy level are affected in the same manner by varying amounts of another element in the arc. Different groups however, may be affected in different manners.

The effect of buffers in eliminating or decreasing the extraneous element effect were investigated. Different amounts of ammonium chloride were found to be ineffective in reducing the effect of added elements. Zinc oxide was found to be very effective in eliminating the effect of lead upon the intensity of the tin lines. In the presence of this buffer, the tin line intensity remained constant within experimental error as the amount of lead was varied.

COMPARISON OF EXPERIMENTAL AND THEORETICAL DATA

The condition selected at which to compare the experimental and predicted values was that in which there were equal numbers of atoms of the test element and of the extraneous element in the sample. The f values were determined for each of the lines of the test elements at this arbitrary level of extraneous element concentration. These were then averaged so that for any given extraneous element and test element combination there was only one f value (Table I).

As the number of test elements is expanded one obtains a large number of lists and the problem of interpretation becomes more complex. To further simplify the set of spectrochemical series shown in the preceding table, the members were all corrected back to the Bi series as a base by multiplying each series by the appropriate factor. When this was done the following arrangement of the elements was obtained:

Na	Sn	Fe	V	Al	Mn
0.52	0.88	0.93	1.01	1.02	1.02
Cd	Bi	Ca	Pb	Zn	Cu
1.04	1.05	1.07	1.12	1.17	1.18

In this array any element will depress the intensity of lines of elements on its right and likewise will enhance the lines of elements on its left. This arrangement is qualitative and the numbers only give a rough idea of the general spacing between the elements. Certain of the characteristics of the effects have been averaged out in this process and it is felt that for a more accurate application of a spectrochemical series the data such as given in Table I are of more value.

A comparison of the predicted series with that obtained experimentally is shown in the Fig. 3. In view of the approximations which were necessary in the theoretical treatment and the magnitude of the error involved in the experimental determinations, the correlation between the two sets of data appears to be good enough to allow the use of this

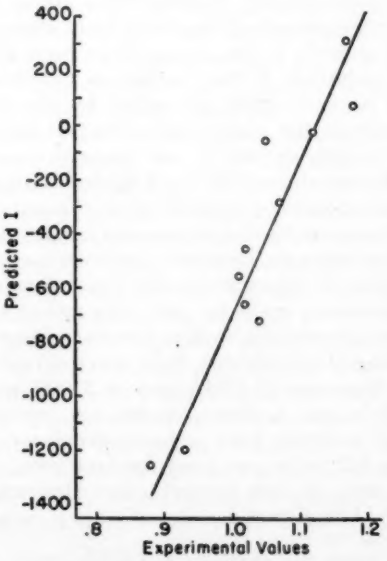

FIG. 3.

method for the prediction of inter-element effects in simple cases. From the straight line relationship displayed in Fig. 3 the proportionality constant of Eq. (2) can be evaluated giving

$$f=1.32-0.284E_i-0.000747Bp_i+0.000178E_iBp_i.$$

SUMMARY

The results of this research are summarized as follows. The effect of the extraneous element varies with the amount of extraneous element present but not in a simple linear fashion. The effect of an extraneous element on a group of lines of an element beginning in the same energy state or ending in the same energy state is the same. However, different groups of lines are affected in different manners. By a study of the experimental results, the elements were arranged in a series showing their inter-element effects in the direct-current arc. This series was found to agree well with the arrangement predicted on the basis of the boiling points and excitation potentials of the elements. A method is given for the prediction of the extraneous element effect on an intensity ratio.

TABLE I. Arrangement of extraneous elements and their f factors in order of their effects on certain test elements compared at equal numbers of atoms of the two elements.

Test element									
Cd		Bi		Zn		Pb		Sn	
Ext. Ele.	f	Ext. Ele.	f	Ext. Ele.	f	Ext. Ele.	f	Ext. Ele.	f
Na	0.51	Fe	0.88	Mn	0.61	Al	0.41	Cd	0.92
V	0.80	Cu	1.00	Cu	0.69	Zn	0.47	Fe	0.99
Pb	0.82	Cd	1.01	Al	0.73	Sn	0.51	Ca	1.04
Sn	0.83	Sn	1.01	Ca	0.79	Fe	0.56	Zn	1.05
Mn	1.05	Ca	1.12	Sn	0.83	Bi	0.66	V	1.06
Fe	1.05	Al	1.13	Bi	0.91	V	0.69	Pb	1.23
Al	1.12	V	1.13	Pb	0.92	Mn	0.71	Al	1.42
Bi	1.19	Zn	1.14	V	0.96	Cu	0.96	Cu	1.46
Ca	1.21	Na	1.33	Cd	1.14				
Cu	1.29	Mn	1.35						
Zn	1.82	Pb	1.37						

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

An Auxiliary Working Standard Equivalent to Magnesium Oxide on a General Electric Spectrophotometer

JOSEPH W. GABEL AND E. I. STEARNS
Research Department, Calco Chemical Division of the American Cyanamid Company, Bound Brook, New Jersey

(Received March 4, 1949)

A method of obtaining spectrophotometric data plotted directly, relative approximately to magnesium oxide yet using glass samples, is described. The method consists of measuring a glass filter in series with the reflectance sample and using an opaque white glass as the reference sample. The transmitting glass is chosen so that over the wave-length range of interest its transmittance is approximately equal to the ratio of the reflectance of the opaque white glass to the reflectance of magnesium oxide.

INTRODUCTION

SPECTROPHOTOMETRIC measurements are made primarily either for colorimetric or analytical interpretations. For colorimetric interpretation, the standard of reference is a magnesium oxide surface by international agreement. In 1931 the International Commission on Illumination adopted a resolution which may be translated as follows: "For the colorimetric measurement of opaque materials, the luminance of the specimens studied ought to be expressed as a function of the luminance of a surface of the oxide of magnesium considered under the same conditions of illumination and observation."[1,2]

For analytical interpretation in which it is sometimes desired to determine spectrophotometrically the identity and quantity of coloring material in a reflectance sample, the theoretical reference is a perfect reflector.[3] While magnesium oxide is not a perfect reflector but has a reflectance of only 99 percent,[4] this is so close to 100 percent that in practice it is used.

Magnesium oxide standards are difficult to prepare and are not stable.[2] The difficulties of preparation are generally admitted, and detailed instructions are presented in letter circular NBS LC-547, March 17, 1939. Vitrolite working standards are issued by the National Bureau of Standards to correct magnesium oxide standards for aging (LC-929).

Spectrophotometric data relative to magnesium oxide for colorimetric work or relative to absolute reflectance for analytical or theoretical work may

[1] Proc. C.I.E., Eighth Session, Cambridge (September 1931), p. 23.
[2] Harry J. Keegan, J. Opt. Soc. Am. **38**, 658 (1948).
[3] O. W. Pineo, U. S. Patent 2,218,357.
[4] Benford, Lloyd, and Schwarz, J. Opt. Soc. Am. **38**, 445–7 (1948).

TABLE I. Reflectance of a Vitrolite glass relative to magnesium oxide and relative to "Compensated Opal."

Wave-length mμ	Reflectance relative to magnesium oxide	Reflectance relative* to "Compensated Opal"	Difference
400	0.913	0.903	+0.010
10	0.903	0.906	−0.003
20	0.898	0.904	−0.006
30	0.897	0.901	−0.004
40	0.896	0.900	−0.004
50	0.902	0.902	0.000
60	0.911	0.909	+0.002
70	0.914	0.912	+0.002
80	0.915	0.914	+0.001
90	0.917	0.914	+0.003
500	0.919	0.917	+0.002
10	0.920	0.917	+0.003
20	0.921	0.918	+0.003
30	0.921	0.918	+0.003
40	0.922	0.919	+0.003
50	0.922	0.919	+0.003
60	0.922	0.918	+0.004
70	0.921	0.917	+0.004
80	0.919	0.915	+0.004
90	0.917	0.913	+0.004
600	0.914	0.912	+0.002
10	0.911	0.908	+0.003
20	0.908	0.905	+0.003
30	0.905	0.902	+0.003
40	0.902	0.900	+0.002
50	0.901	0.898	+0.003
60	0.899	0.896	+0.003
70	0.898	0.894	+0.004
80	0.897	0.892	+0.005
90	0.896	0.889	+0.007
700	0.895	0.887	+0.008
10	0.893	0.882	+0.011
20	0.891	0.880	+0.011
30	0.889	0.878	+0.011
40	0.887	0.875	+0.012
50	0.885	0.872	+0.013

* Note: These values were obtained by using two clear glass filters in series in the sample beam and one clear glass filter in the reference beam.

be calculated from data obtained instrumentally relative to a working standard from the National Bureau of Standards. However necessary, this is time consuming and laborious and it would be preferable to have the instrument plot the desired data directly.

"COMPENSATED OPAL" METHOD

A method of measurement employed in this laboratory has the advantage of giving data relative approximately to magnesium oxide without requiring the daily preparation of magnesium oxide. With a General Electric type spectrophotometer,[5,6] a white opaque glass, which has been called an opal glass standard, is used on the reference side of the spectrophotometer and included in the sample beam is a clear glass filter. This combination has been called "Compensated Opal." The locations of the filter and the opal block are shown in Fig. 1. This filter is fixed in a holder sliding in a set of ways located directly in front of the entrance port of the integrating sphere. A description of this filter holder has been published previously.[7]

When it is desired to measure transmission samples, the glass filter is easily slid out of its position in the sample beam, leaving in its place a blank aperture of the filter slide. A second piece of opal glass equivalent to the reference standard of the "Compensated Opal" combination is then placed at the reflectance port of the sample beam for transmission measurements.

MATHEMATICAL CONSIDERATION

In a Martens' photometer used for measurement of reflectance samples, the radiant flux in the reference beam is $R_mP\sin^2\alpha$, and the flux in the sample beam is $R_sP\cos^2\alpha$ where: P = radiant flux leaving the photometering Rochon prism, α = angle of the photometering Rochon prism relative to the Wollaston prism, R_m = reflectance of the magnesium oxide standard, and R_s = reflectance of the sample.

In the normal operation, the photometering Rochon prism is adjusted until the flux in the two beams is equal. Then:

$$R_mP\sin^2\alpha = R_sP\cos^2\alpha$$

or

$$R_s = R_m\tan^2\alpha.$$

When measurements are made with the "Compensated Opal," the flux in the reference beam is $R_0P\sin^2\alpha$ and the flux in the sample beam is $T_gR_sP\cos^2\alpha$ where the additional symbols are R_0 = reflectance of the opal and T_g = transmittance of the glass. When these two beams of flux are made equal, $R_s = (R_0/T_g)\tan^2\alpha$.

Now, if $(R_0/T_g) = R_m$ at each wave-length, the measurement of R_s with "Compensated Opal" is exactly equivalent to the measurement of R_s relative to magnesium oxide. This is accomplished if the transmittance of the glass equals the reflectance of the opal relative to magnesium oxide, or $T_g = R_0/R_m$.

CHOICE OF COMPONENTS

An opal glass should be chosen which has a smooth surface so that it can be easily cleaned. The glass plate should be highly polished so that it also can be cleaned easily. Both the opal and the glass should be large enough to accommodate the beam satisfactorily. Ideally, a specific glass sample should be chosen which has a transmittance approximately corresponding to the reflectance of the opal glass measured against magnesium oxide.

If the correspondence of T_g with R_0/R_m is not within the desired limits, several special techniques may be employed. A thick piece of glass with low

[5] O. W. Pineo, U. S. Patent 2,107,836.
[6] Arthur C. Hardy, J. Opt. Soc. Am. **25**. 305–311 (1935).
[7] R. H. Kienle and E. I. Stearns, Instruments **20**, 1057–64 (1947).

T_g values may be ground thinner. A glass plate with high T_g values may be mounted at an angle to the light beam to reduce T_g. By mounting one glass plate in the reference beam and another of slightly different color in the sample beam, T_g may be increased at some wave-lengths and decreased at others.

In order to avoid an extensive search for a single glass filter with satisfactory characteristics, an extra pair of clear glass plates has been used, one in each beam.

One of the white structural Vitrolite glass working standards issued by the National Bureau of Standards (NBS Test 116338) was measured with the combination of an opal block plus a glass plate in the reference beam and two glass plates in the sample beam. The calibration data of this Vitrolite standard relative to magnesium oxide obtained from the National Bureau of Standards are tabulated in column 2 of Table I. Column 3 lists the observed values when measured against this "Compensated Opal." It may be noted that the data obtained by measuring against this "Compensated Opal" agree within 0.5 percent with the reflectances as measured at the National Bureau of Standards against a magnesium oxide over the wave-length range of 430–680 mμ, which is within the limits of reproducibility of magnesium oxide surfaces.[2]

This "Compensated Opal" was further checked by measuring a magnesium oxide surface as a sample and the differences from 100 percent were found to be about the same as the differences listed in column 4 of Table I.

DISCUSSION OF THE METHOD

As a variation, it is possible to plot relative to a completely reflecting standard instead of magnesium oxide if it is so desired in analytical or theoretical work. This would be accomplished by using a slightly more absorbing piece of glass in the sample beam. The method of a compensated standard could probably be applied to other optical instruments than this spectrophotometer if it were so desired.

One possible error which might be introduced in this method is to have the glass filter inserted in the beam in a non-reproducible position. The transmittance of a transparent piece of glass, of course, varies with the angle of incidence of the light. This error has been circumvented by mounting the glass in the filter slide which is very reproducibly placed in the beam. Another possible source of error which also can be largely eliminated is back-reflectance from the glass sample similar to that which has been reported in the measurement of an interference filter.[3]

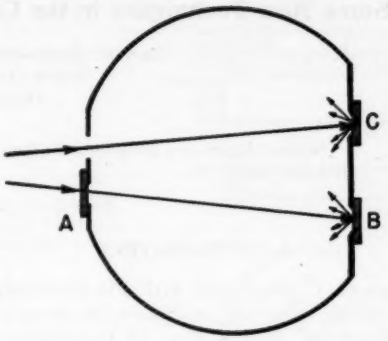

FIG. 1. A schematic diagram of the integrating sphere showing the location of the glass filter. A—The glass filter in the sample beam. B—Sample. C—Opal standard (the combination of A and C comprises the "Compensated Opal").

A third possible source of error is a variation in the amount of specular light included in the measurement. This is circumvented in the G.E. spectrophotometer by using 6° incident light.

A fourth possible error, aging of the opal glass or transmitting glass, can be circumvented by an occasional check measurement of MgO.

It is necessary to smoke the magnesium oxide standard regularly in order to get reproducible data, and it is necessary to clean the "Compensated Opal" occasionally to get reproducible data. The advantage of the "Compensated Opal" method is that it is easier to clean its smooth glass components than it is to smoke a MgO standard. A second advantage of the "Compensated Opal" method is that absolute reflectance data can be plotted as easily and accurately as reflectances relative to magnesium oxide by making $Tg = R_0$.

---

[3] W. E. K. Middleton, J. Opt. Soc. Am. **38**, 74 (1948).

Some New Techniques in the Construction of Spectrographic Absorption Cells

GILBERT KIVENSON, ALBERT ROTH, AND MYRTLE RIDER
Mellon Institute, Pittsburgh, Pennsylvania
(Received February 24, 1949)

Several designs are presented for the construction of pressure absorption cells for use in the infra-red and ultraviolet.

A. INTRODUCTION

IN a project* concerned with the development of automatic infra-red analyzers for recording and controlling the composition of butadiene-styrene mixtures, it became necessary to design an absorption cell which would continuously withstand pressures of 100 to 125 pounds per square inch. Since radiation of wave-lengths greater than 3.5 microns was not required for the analysis, the use of quartz was permitted as the window material. The desired spacing was approximately 1.5 millimeters. The absolute value of the spacing was not as important, however, as the constancy which could be obtained under relatively high pressures.

B. PRELIMINARY WORK

A number of liquid cell designs have been described in the literature.[1–6] These in general utilize a packing material (such as rubber or lead) and depend on externally applied pressure to maintain the seal. Other methods rely on a cement of some type. It was found, however, that most of the cements described are attacked by styrene or do not bond well to hard, non-porous surfaces. Several cells using a packing were built, but these failed after several days service, usually owing to swelling of the packing or to uneven distortion of the spacing material.

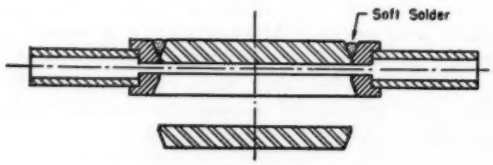

FIG. 1. Cross section of soldered construction quartz absorption cell.

The direct fusion of quartz to metal by means of graded seals was considered, but the difficulties of annealing (especially in larger cell sizes) made this procedure seem impractical. A lower temperature method involving platinizing of the quartz and direct soldering of the quartz to metal was then tried. This process, combined with a suitable mechanical design, proved satisfactory in the cell construction.

C. CONSTRUCTIONAL DETAILS

Figure 1 shows the most simple form of the soldered construction cell. The cell frame can be made of brass and silver-plated prior to the final assembly. A chamfer is cut into both faces, allowing a small perpendicular edge to remain in the middle of the frame. The width of this surface is governed by the desired cell spacing. The inlet and outlet holes are then bored, an enlargement being cut at the end of either to permit lead tubing to be silver-soldered into place. A bevel is next ground on each quartz disk. It is important that this bevel be finely polished to prevent air holes in the later stages of the process. Five to seven coats of platinum solution[7] are fired on the bevels of the disks. The platinum is then plated with copper and lightly tinned, using soft solder and a rosin flux. The copper serves to protect the platinum surface from being dissolved by molten solder. Excess platinum coating is removed from the surface of the quartz by the use of the pitch lap and a water slurry of cerium oxide as the polishing material.

A bent wire constituting the spacer is inserted before the final soldering. With the disks in place, soft solder is made to flow into the "V" groove by applying a soldering iron to the edge of the frame and "puddling" the molten solder until a uniform layer is obtained. The heat is removed and the solder allowed to solidify. The process is then repeated on the other side of the cell. The spacer wire is removed and suitable valves or couplings added.

These cells were tested by the application of nitrogen pressure while immersed in water. Bursting strength for the smaller cells was greater than 150 pounds per square inch.

* This development is part of the program sponsored by the Office of Rubber Reserve, Reconstruction Finance Corporation, and has been approved by that office for publication.

[1] W. W. Coblentz, Publication No. 35, Carnegie Institute of Washington (1905).
[2] W. West and R. T. Edwards, J. Chem. Phys. 5, 14 (1937).
[3] C. Corin and G. B. B. M. Sutherland, Proc. Roy. Soc. A165, 43 (1938).
[4] H. M. Randal, Rev. Sci. Inst. 10, 195 (1939).
[5] E. S. Barr, Rev. Sci. Inst. 12, 396 (1941).
[6] L. Gildart, and N. Wright, Rev. Sci. Inst. 12, 204 (1941).
[7] E. C. McHelvy and C. S. Taylor, J. Am. Chem. Soc. 42, 1364 (1920).

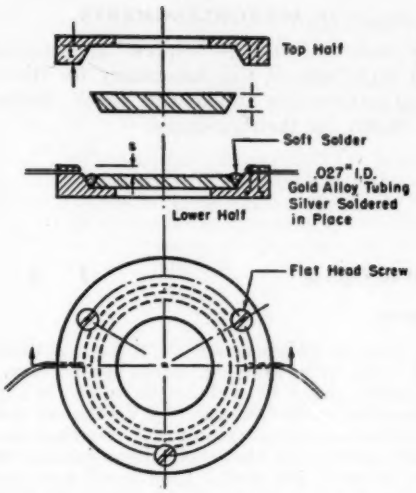

FIG. 2. Improved mechanical form, bolted plate type.

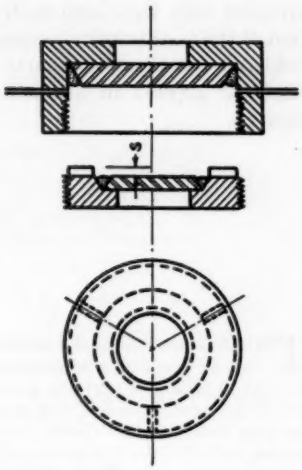

FIG. 3. Improved mechanical form, threaded type.

D. IMPROVED DESIGNS

The cell described, although simple to construct, has the basic structural difficulty that the soldered joint is both the sealing and strength feature. It was felt that long-range reliability could be better achieved if additional mechanical reinforcement could be provided. On the basis of suggestions made to them, the authors therefore designed and constructed two improved models, illustrated in Figs. 2 and 3. In both cases the major portion of the tensile load is born by mechanical components and the solder serves only as the seal. In addition, the reversal of the "V" grooves causes a greater compacting of the seal under pressure. The initial assembly is much the same as the older type except that a hot plate can be used for soldering the quartz disks in place. Eighth-inch flexible tubing (gold alloy for inertness) is used for entrance and exit leads. Final assembly of the two halves (Fig. 2) is made with a low melting alloy "Cerroseal"** so that the soft solder joints are not disturbed. In the type shown in Fig. 3*** the threads are tinned with "Cerroseal" before assembly and the final threading is done on a hot plate controlled to the melting temperature of the alloy. The threaded type has an advantage in that it can easily be taken apart for cleaning, without disturbing the quartz-to-metal soldered joints, by melting the "Cerroseal" and unscrewing the inner section. In both types, the spacing "*S*" is set by the initial machining.

In view of the more complicated assembly process of the improved cells, the earlier model was used for all laboratory work and for experiments in which continuous pressures were not applied for periods longer than one week. The improved models have been employed in a wider variety of analyses, including continuous plant stream sampling at average pressures of 115 pounds per square inch for periods of six months. No failures have occurred in these applications.

E. EXTENSION OF THE PROCESS

Experiments have been conducted to determine whether the general method could be applied to other infra-red transmitting materials.

A means of using silver chloride was found. It consists of vacuum-evaporation of a thick layer of silver on the beveled edge of a silver chloride**** disk (cut in a manner similar to the quartz disks). The adherence of the evaporated film to the silver chloride facilitates excellent bonding when the disk is soldered into place. It was found necessary to evaporate the silver on a portion of the flat periphery of the disk as well as the beveled edge in order to prevent solder from coming in direct contact with the silver chloride. Cell frames were of the type presented in Fig. 2. Silver chloride cells were tested at 50 pounds per square inch, but, as is well known, the cold flow of this material determines the upper working pressure for a given application.

F. CONCLUSIONS

A process has been developed for the construction of reliable high pressure absorption cells for use in the near infra-red and ultraviolet. Medium

** Cerroseal is supplied by the Cerrode Pasco Copper Corporation, 40 Wall Street, New York 5, N. Y.

*** This type of cell is being manufactured by the Buhl Optical Company, 1009 Beech Avenue, Pittsburgh 12, Pa.

**** Harshaw Chemical Company, Cleveland, Ohio.

pressure cells were also built that enable application of the transmitting properties of silver chloride. Soldered construction quartz cells have been successfully applied in automatic analyses of plant streams.

G. ACKNOWLEDGMENTS

The authors wish to express appreciation to Ralph Steinback of this laboratory for his active aid and cooperation and to Drs. O. W. Burke and E. E. Stahly for their guidance.

International Commission of Optics

Diffraction Grating Survey

THE Swedish National Committee has been given the task of conducting a world-wide survey on diffraction gratings to determine: (a) the total need for such gratings; and (b) the actual possibility of obtaining them. This survey has been under way for about a year, and is quite well advanced insofar as the European countries are concerned.

The appropriate members of the Swedish Committee have been in touch with present American manufacturers of diffraction gratings and also with the several new laboratories which are undertaking their manufacture. However, the Committee has not yet been able to obtain the answer to the first question listed above, regarding the total need for diffraction gratings in the U. S. A. Although it would be felt by most American spectroscopists that this is a subject which, if not already under fair control, will soon be adequately handled within the U. S. A., still it seems highly desirable to cooperate with the International Commission of Optics in making this important survey.

In view of the serious practical difficulties of identifying and circularizing all persons in the United States who might be interested in obtaining diffraction gratings, it has been decided to run this notice as the best practical way of reaching most or all of these people. It is therefore requested that all persons concerned communicate their answers to the questions listed below either directly to Dr. E. Ingelstam, Department of Physics, Royal Institute of Technology, Stockholm 26, Sweden, or to the Chairman of the U. S. National Committee of the I. C. O., whose name and address appear below.

The specific points which should be covered in any such communication are the following: (It is understood, of course, that your answers are only estimates, and in no sense represent commitments.) (1) Number of gratings probably needed; (2) Size of ruled area desired (give length and breadth); (3) Radius of curvature desired; (4) Line spacing desired; (5) Material (at present aluminized glass seems to be the standard surface: if other material is required, this should be stated); (6) Special qualities, such as concentration of intensity in a certain order or special requirements on resolving power; (7) Would replica gratings be satisfactory for your needs? The methods of making replica gratings have advanced considerably and it may soon be possible to obtain rather inexpensive replica concave gratings which, although not equal in quality to original gratings, are good enough for many purposes.

Replies to this questionnaire by all interested persons in the U. S. A. would be greatly appreciated by the Swedish National Committee and would significantly increase the value of this important survey. Your cooperation is earnestly solicited.

STANLEY S. BALLARD,
Chairman, U.S.N.C.-I.C.O.,
Tufts College,
Medford 55, Massachusetts

Oblique Reflecting Systems

BERNARD S. LEE
Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey
(Received March 29, 1949)

The system of mirrors which will deviate a ray by an angle D while rotating an image an angle R is such that the dihedral angle A between the two mirrors is in accord with $\cos A = \pm\cos D/2 \cos R/2$. The position of the common edge bears a fixed but simple relationship to the initial and terminal ray. The system of two mirrors with fixed dihedral angle may be rotated about the common edge of the two mirrors as an axis to provide, within limits, a continuous multiplicity of solutions. Design techniques are included.

INTRODUCTION

DESIGNERS of optical systems such as are found in film viewing devices, television projectors, data recording apparatus, etc. are frequently confronted with the utilization of plane reflecting surfaces. Systems which lie in one plane or in mutually perpendicular planes are relatively easy to comprehend; whereas oblique reflection systems are not so simple. A ray may be re-directed in any desired new direction by a single reflection. By means of two reflections, both re-direction and rotation of an image plane normal to the ray (the field) may be accomplished. For this reason, oblique two-mirror systems may be especially useful. For any given requirement of re-direction and rotation a multiplicity of solutions exist. The designer's problem, therefore, becomes a matter of selecting the most desirable of the group of possible answers.

DESIGN TECHNIQUES

A useful technique, familiar to many designers of optical systems, is that of laying out on paper the optical components of a system in a straight line; then by cutting along the outside rays of a typical cross section the layout may be removed from the paper. The effect of introducing plane reflectors into the system may be studied by folding this cut-out longitudinal section, each fold representing the intersection of the optical beam with a reflector. The relative position of components and images may be observed and design features such as self obstruction of light beam by components may be noted. By trial and error a happy arrangement of mirrors, parts, images, and apertures may be found. The analogy of the strip of paper and the light beam is valid provided that the paper is neither stretched nor compressed and provided that the sections between folds representing reflections remain planes. The technique of "folding" an optical system may be applied to three-dimensional as well as coplanar systems. A complete three-dimensional model of a light beam may be constructed by folding the strip of paper lengthwise to form a right angle. The two members at right angles represent two characteristic longitudinal sections of the beam and this folded strip may then be folded in two planes to represent intersections with plane reflectors.

THE PROBLEM

Given two directed lines in space representing the original and terminal axial rays of an image forming system, also given an original image orientation and a desired final image orientation, how shall two reflectors be placed in order to achieve the desired terminal direction and orientation?

GEOMETRIC SOLUTION

The geometric solution (Fig. 1) requires a special set of x, y, and z coordinates. Let the normal to the common plane of the initial and terminal ray be the z axis. Let the projection of the terminal ray on this common plane be the y axis. The x axis is then the normal to the y axis in the common plane of the initial and terminal rays. The angle made by the initial (original) ray with the y axis is D, the angle of deviation. If this deviation were accomplished by a *single* reflection, the image plane (field) would revolve with the ray about a normal

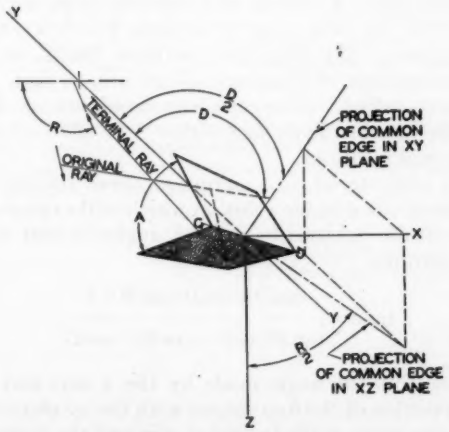

FIG. 1. Geometric solution.

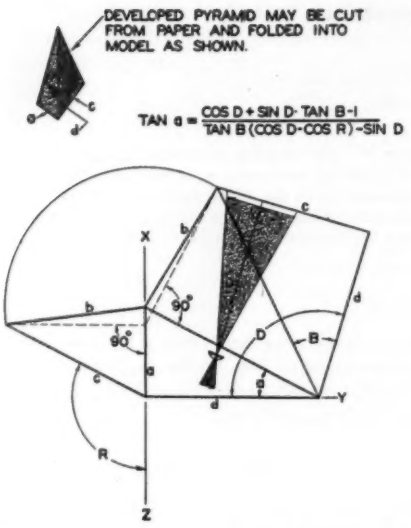

FIG. 2. Developed pyramid of extended light beam segments.

to the incident and reflected ray at the point of reflection. This condition of image orientation, achieved by a single reflection, we shall call *zero* rotation. When re-directed and re-oriented by a two mirror system, we shall refer to the angular change from this condition as the angle of rotation R.

The requirement for any value of D and R is that the two mirrors shall form a dihedral angle A such that

$$\cos A = \pm \cos D/2 \cdot \cos R/2,$$

and the common edge of the two mirrors shall project in the xy plane at an angle of $D/2$ with the y axis and in the xz plane at an angle of $R/2$ with the z axis. The above conditions describe a system of mirrors which may be rotated (with fixed dihedral angle A) about the common edge without affecting the redirection or rotation. For this reason, it appears that there are, between limits, an infinite number of solutions. There are, in fact, two systems which will meet given requirements. Due to the plus or minus sign, either A or $180° - A$ provide real solutions.

In order to select a unique solution, reference is made to the graphic solution which will explain the use of the arbitrarily defined angles a and B in the formula

$$\tan a = \frac{\cos D + \sin D \cdot \tan B - 1}{\tan B(\cos D - \cos R) - \sin D},$$

where B is the angle made by the y axis and the intersection of the first mirror with the xy plane and a is the angle made by the y axis and the intersection of the yz plane. Fixing either a or B will provide the unique solution required for design purposes.

GRAPHIC SOLUTION

A useful concept derived from the paper-folding and model making techniques previously mentioned is that the three folded segments of a twice reflected system may be extended to form a pyramid having three sides and a base. This pyramid when developed as in Fig. 2 presents a convenient form for the study of a beam of light and the manner in which it intersects reflecting surfaces.

Each corner of the pyramid represents a reflection or "fold." Unfavorable angles of reflection, abnormally large reflecting surfaces, most economical location of reflectors where they are to be a surface of a glass prism, direction and orientation of images, etc. may be noted on this graphical presentation.

Figure 3 illustrates the manner in which a and B vary as the pair of mirrors with fixed dihedral angle is rotated about the common edge.

Figure 4 shows a systematized construction of a three dimensional model which may be useful for examining the course of rays and position of parts in any characteristic cross section. This illustration shows a section in the plane of the paper and another section revolved 90° from the normal to the paper. Any section may be regarded as a portion of a pyramid as in Fig. 2. A useful fact in the construction of these models is that the locus of apices of all such pyramids will be the common edge of the two reflecting surfaces.

VECTOR ANALYSIS

L. Silberstein[1] published in 1916 the vector notation for dealing with reflections from any number of plane mirrors. A useful application of his analysis was to provide a formal proof that the cube corner is the only solution to the requirements for a "central" mirror. E. Kaspar[2] 1948 applies vector analysis to both reflecting and refracting systems.

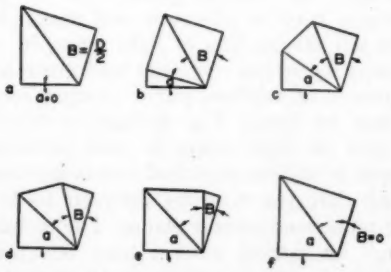

FIG. 3. Effect of rotating mirror system around the common edge.

[1] L. Silberstein, Phil. Mag. **30–31**, 487–494 (1916).
[2] E. Kaspar, Revue D'Optique, 15–28 (January 1948).

Neither of these papers deals especially with the orientation of the image as defined herein nor the problem where the direction of rays and image orientation are given and the mirror positions are unknown. It is possible to extend their methods to provide this information, however.

SPECIAL CASES

Certain cases are of special interest. For instance, in the formula $\cos A = \pm\cos D/2 \cos R/2$, when $A = 0°$ the mirrors are parallel and neither deviation nor rotation can occur. When $R = 0$ the formula is identical with that for deviation by two mirrors in a coplanar system of reflections. When $A = 90°$ either deviation or rotation must be 0° or 180°. This is the case of the roof prism as may be demonstrated by looking into the ridge of the roof and rotating the prism on the line of sight. When either a or B of the formula of Fig. 2 is reduced to zero, the rays of the system lie in mutually perpendicular planes. This is the case of the Porro prism and may also be found applied in certain alt-azimuth instruments.

CONCLUSIONS

For design purposes, the trigonometric formulas and the graphic solutions offer a practical approach to the problem of oblique reflecting systems. In order to deal with the general case without resorting to a special coordinate system, the use of vector analysis is indicated.

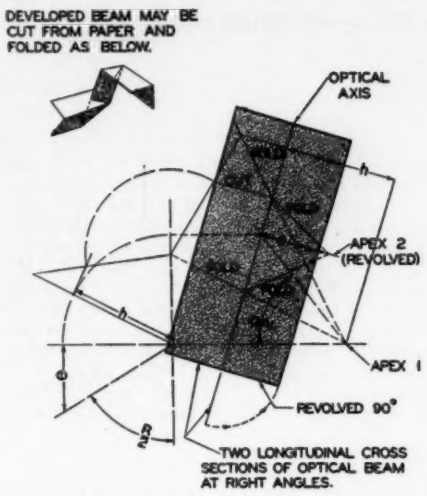
DEVELOPED BEAM MAY BE CUT FROM PAPER AND FOLDED AS BELOW.
OPTICAL AXIS
APEX 2 (REVOLVED)
APEX 1
REVOLVED 90°
TWO LONGITUDINAL CROSS SECTIONS OF OPTICAL BEAM AT RIGHT ANGLES.

FIG. 4. Construction of three-dimensional model.

Book Review

The High Current Carbon Arc

WOLFGANG FINKELNBURG. Final Fiat Report 1052. Office of Military Government for Germany (U.S.). Pp. 219, figs. 129.

The book is an excellent and complete report on the work done in Germany, chiefly by the author and his group, on one important form of the carbon arc, called the *high current* arc. Its publication will be welcomed not only by men in industrial laboratories, for whom the report should be of great value, but also by more academically minded physicists who desire an understanding of the principles underlying the curious phenomena to which the arc form of a discharge gives rise.

A good deal was known about the technical aspects of the so-called Beck arc, a type of discharge that has been studied for its utility ever since its discovery in 1910. But the knowledge available was of a phenomenological sort, uncoordinated in a basic sense and always looking to applications. It has been Finkelnburg's concern to tie together, by significant experiments and happy conjectures, the utilitarian bits of information into what promises soon to be a satisfactory theory of *high current* discharges.

New points of view and new explanations are found throughout the book. The Beck arc is shown to be a special form of a more general kind of discharge, here called the *high current* carbon arc. An explanation is given of the mechanism responsible for the hissing of over-charged solid-carbon arcs, for the radiating anode flame and numerous other interesting observations. More than 30 pages and many diagrams and tables are devoted to the optical and radiative properties of the arc, and a good deal of attention is paid to the temperature of the discharge under various operating conditions. The last chapter of the book, dealing with technical applications (sources for search lights, studio illumination, movie projection, therapeutics, chemical applications) will have a special appeal to industrial researchers.

The publication of this book is fortunate; credit is due both the author for pioneering in an important field and the military government for its good judgment in allowing this record to be published.

HENRY MARGENAU,
Yale University, New Haven

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

The Herschel Effect and the Structure and Stability of the Photographic Latent Image*

G. KORNFELD
Kodak Research Laboratories, Rochester, New York
(Received April 4, 1949)

In a previous investigation carried out with a contrast motion-picture positive emulsion of uniform grain size, it was found that for the first few minutes of development in a standard Elon-hydroquinone developer the differential density decrease produced by the Herschel effect was constant in each exposure step of the pre-exposure to white light. This led to the conclusion that the "critical size" of the development centers remained essentially unchanged during that time.

In the present investigation with the same emulsion, the constancy of the Herschel effect was confirmed, both with the same developer and with a pure surface ferrous oxalate developer. When development baths were used which contained varying amounts of sodium thiosulfate, the Herschel effect decreased, corresponding to the solvent action. This agrees with the widely accepted view that the Herschel effect produces a redistribution of centers through the grain.

The Herschel effect proved also to be a sensitive indicator of the growth of surface centers which occurred when white-light and infra-red exposures were separated by longer time intervals. When the film was kept at higher temperatures, the decrease of the Herschel effect was the only evidence of latent-image change. Furthermore, the very great difference in the Herschel effect corresponding to different intensity levels of pre-exposure gave a striking illustration of the difference in center size produced at these different intensity levels.

I. HERSCHEL EFFECT AND DEVELOPMENT TIME IN SURFACE DEVELOPERS

IN a previous investigation[1] it was shown that the Herschel effect can be treated as a first-order reaction if it is assumed that the absorption of infra-red light is proportional to the amount of latent image and that the developed silver is proportional to the latent image. The proportionality between density and developed silver was tested experimentally and found to be sufficient for the purpose. It was further considered likely that at low and medium exposures very few grains would contain more than one development center or a development center of more than minimum size, so that the loss of one silver atom from a center would nearly always result in the loss of one developable grain. For the particular emulsion used, it could be assumed that the probability of containing a center was equal for all the grains, since there was little deviation from an average size and shape among its grains. The differential density decrease produced by the Herschel effect was supposed, therefore, to be proportional to the original density, and $\log a/a-x$ was taken as a measure for the reaction speed, where a signifies the original density, and x the density decrease. $\text{Log}a/a-x$ was found to be proportional to the time of infra-red exposure from the very beginning, and this lack of an induction period justified the assumption that a center needs only one absorbed quantum of infra-red light to exhibit the Herschel effect. Unless it is supposed that the center underwent complete destruction in this photochemical reaction—and there are various reasons against such a supposition—the loss of one silver atom must be responsible for the loss of developability of one grain.

$\log a/a-x$ was found to be constant for the same exposure step from 2 to 5 minutes' development time, up to which time the density curves for different development times were found to be converging. From this constancy it was concluded that the same number of grains were developed for each exposure step from 2 to 5 minutes' development time, and that the density increase between 2 and 5 minutes was produced by the progress of development inside the grains. With longer development times, which correspond to a parallel shift of the density curves, the values of $\log a/a-x$ decreased, indicating an increase in the number of developed grains; it must be assumed that beyond 5 or 6 minutes, the developer could react with grains whose latent-image specks were below the critical size.

The span of time for which a definite critical size is required for the development of the latent image will vary with the emulsion and the developer used. Evidence for "explosive" development was found with some rapid developers, which means that, there, any density increase during measurable times must be attributed to the development of grains with smaller specks. It seems, however, that for many emulsions and developers a finite time interval will be taken up by the progress of development inside the grains which have a latent-image speck at or above the critical size. This general conclusion must be drawn from the investigations of Burton and Berg on image and subimage[2] and on the conditions under which the subimage can be made developable.

Kodak D-16 Developer, the Elon-hydroquinone developer used in the previous investigation, had the effect of a surface developer, although it con-

* Communication No. 1251 from the Kodak Research Laboratories.

[1] G. Kornfeld and T. H. James, J. Opt. Soc. Am. **33**, 615 (1943).

[2] P. C. Burton and W. F. Berg, Phot. J. **86B**, 2 (1946); P. C. Burton, Phot. J. **86B**, 62 (1946); W. F. Berg, Rep. Prog. Physics **XI**, 248 (1948).

TABLE I. Developer D-16, average values of $\log a/a-x$ for various development times.

Exposure step	2 min.	2½ min.	3 min.	4 min.	5 min.	6 min.	15 min.
14	0.19	0.22	0.19	0.27	0.23	0.23	0.21
13	0.30	0.27	0.35	0.29	0.37	0.35	0.19
12	0.33	0.35	0.39	0.35	0.28	0.26	0.21
11	0.27	0.28	0.31	0.27	0.24	0.22	0.17
10	0.23	0.22	0.22	0.19	0.18	0.17	0.11
9	0.11	0.19	0.16	0.13	0.13	0.11	0.09

TABLE II. Ferrous oxalate developer, average values of $\log a/a-x$ for various development times.

Exposure step	3 min.	5 min.	7 min.	9 min.	11 min.
14	0.24	0.25	0.33	0.22	0.25
13	0.27	0.22	0.26	0.16	0.21
12	0.24	0.22	0.23	0.14	0.18
11	0.16	0.15	0.17	0.12	0.13
10	0.11	0.12	0.13	0.09	0.09
9	0.07	0.08	0.09	0.07	0.07

tained some sulfite.** It seemed desirable to test the results with a pure surface developer. The experimental arrangements were unchanged; the filter combination of two Wratten No. 87 filters and one Wratten No. 25 filter was again used for the infra-red exposure. This filter combination, which was first proposed by Webb and Evans,[3] cuts the visible light off at about 750 mμ and eliminates actinic light. This fact is extremely valuable for quantitative measurements and outweighs the disadvantage of the loss of effective spectral regions. The maximum of spectral efficiency for the Herschel effect seems to be quite variable. Carroll and Kretchman[4] found 650 mμ as the center of the optimum region, which, however, was quite broad. Bartelt and Klug[5] found the maximum at 750 mμ, with a shift to longer wave-lengths after continued irradiation. Belliot[6] found the effective region to be between 550 and 770 mμ for pure silver chloride and between 650 and 1100 mμ for pure silver bromide. Gorokhovski and Shestakov[7] stated that the maximum and the short wave-length limits of the Herschel effect depended upon the pBr in the emulsion:

For a pBr of *1.07*, the maximum was beyond $\lambda = 620$ mμ.
For a pBr of *1.30*, the maximum was at $\lambda = 775$ mμ.
For a pBr of *2.12*, the maximum was at $\lambda = 810$ mμ.

The findings of the previous investigation with the Elon-hydroquinone developer were confirmed, although the sensitivity of this emulsion to the Herschel effect turned out to be smaller than that of the previous sample. Accordingly, the opal glass in the Model X sensitometer was replaced by a much more transparent ground glass in order to provide more illumination; this change of intensity should in no way affect the results, since there is no reciprocity-law failure for the Herschel effect. Table I shows, indeed, that $\log a/a-x$ was fairly constant for development times between 2 and 6 minutes and decreased at longer times. Analogous results were obtained with a ferrous oxalate developer,*** as shown in Table II. Log $a/a-x$ was fairly constant between 3 and 7 minutes and decreased for longer development times. Again it was found that the constancy of $\log a/a-x$ held for times where the development curves converged, whereas its decrease coincided with a parallel shift of the density curves. Since there can be no suspicion of any solvent action with this developer, the proposed interpretation gains added weight. New grains are added for development only at longer development times, where the density curves show a parallel shift.

The straight-line portions of single density curves, with and without infra-red exposure, were found to be parallel, as shown in Fig. 1. The greatest relative density decrease was therefore at the beginning of the straight-line part. This was found to be a characteristic for surface development in this emulsion. It was not a general characteristic, however,

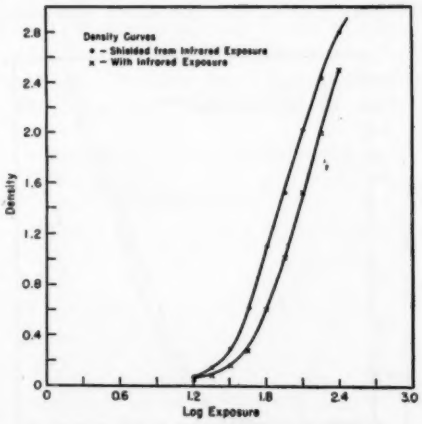

FIG. 1. Contrast motion-picture positive emulsion used throughout the investigation. Development in D-16—4 min. at 68°F.

**.A one-liter development bath contains: 0.31 gram of Elon; 6.0 grams of hydroquinone; 39.6 grams of sodium sulfite (desiccated); 18.7 grams of sodium carbonate (desiccated); 0.86 gram of potassium bromide; 0.86 gram of citric acid; and 1.5 grams of metabisulfite.

[3] J. H. Webb and C. H. Evans, J. Opt. Soc. Am. **28**, 249 (1938).

[4] B. H. Carroll and C. M. Kretchman, J. Research Nat. Bur. Stand. **10**, 449 (1933).

[5] O. Bartelt and H. Klug, Zeits. f. Physik **89**, 779 (1934).

[6] H. Belliot, Travaux et Mémoires du L. C. I. M., Cahier 1, 47 (1945).

[7] Y. N. Gorokhovski and S. A. Shestakov, J. Phys. Chem. U.S.S.R. **11**, 356 (1938).

*** This developer comprises 1000 grams of potassium oxalate in 4 liters of water mixed with a solution of 330 grams of ferrous sulfate in 1 liter of water, to which 1 cc of concentrated sulfuric acid had been added.

BOSTON UNIVERSITY LIBRARIES

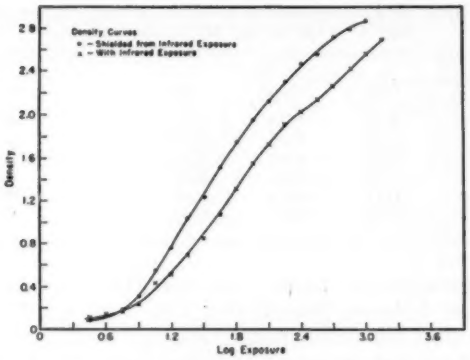

FIG. 2. Another commercial emulsion. Development in D-16—4 min. at 68°F.

as Falla[8] claimed. Figure 2 shows density curves, before and after infra-red exposure, for another emulsion where the density decrease increased at higher densities and the relative density decrease was fairly constant over many exposure steps.

II. HERSCHEL EFFECT WITH A SOLVENT DEVELOPMENT BATH

The Herschel effect was further investigated in the solvent developer, D-16, containing various amounts of thiosulfate. An unexposed strip was given the same treatment in the development bath to provide a sample for analysis of the dissolved silver bromide. With increasing amounts of thiosulfate, the density curves changed slope according to the change from chemical to physical development. The Herschel effect still showed its optimum at the beginning of the straight-line part of the characteristic curve, but the effect was decreased even there and showed a much stronger decrease at higher densities. Higher thiosulfate content and longer development times practically eliminated the effect in the overexposure region; occasionally, the effect was even reversed, but the reversal was within the limit of experimental error, never exceeding a density difference of 0.02. This absence of a "positive effect" below the solarization region is in agreement with the findings of Debot, Falla, and Hautot,[9] who investigated the redistribution of centers between the surface and the interior of the grain under the influence of infra-red radiation.

The Herschel effect was tested at various thiosulfate concentrations and development times, beginning with 0.5 percent thiosulfate concentration and 3-minute development time and extending to a thiosulfate concentration of 1.5 percent with 14-minute development time. These tests correspond to the whole range of solvent action, determined on the comparison strips, from the loss of 19 percent until practically complete fixing out. The values determined may be only lower limits for the exposed strips, since the solvent action on the unexposed strips should be less intense than that for the developed grains.[10] The depth of penetration of the solvent developer for the unexposed grains can, however, be calculated from these values. This could be done with the emulsion used, since its grains deviated little from an average size and spherical shape. Table III lists the experimental conditions, together with the corresponding values for the remaining percentage of silver bromide and the values for the depth of penetration, Δr. This depth was, accordingly, 26 mμ for the action of 0.5 percent thiosulfate during 3 minutes. This result contradicts Debot's findings.[11] From data extrapolated from the literature, he calculated the

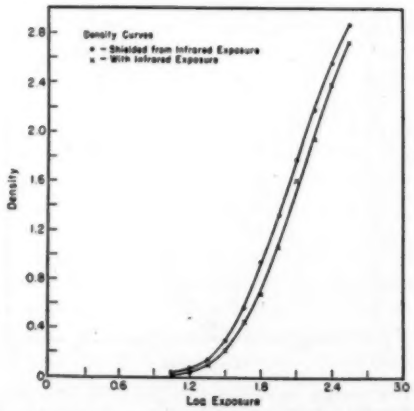

FIG. 3. Density curves after keeping eleven days between white and infra-red exposures. Development in D-16—4 min. at 68°F.

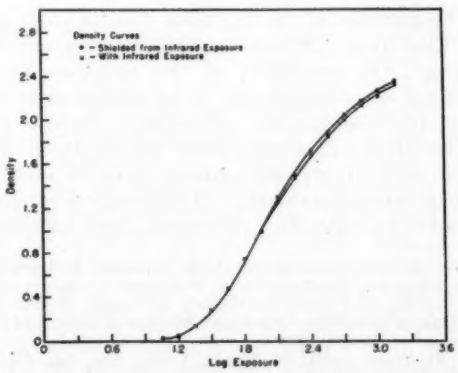

FIG. 4. Density curves after keeping three months between white and infra-red exposures. Development in D-16—4 min. at 68°F.

[8] Cf. A. Hautot and L. Falla, Bull. Soc. Roy. Sci. Liège 11, 298 (1942).

[9] A. Hautot, Bull. Soc. Roy. Sci. Liège 14, 200 (1945).
[10] H. A. Miller, *PSA Journal* 14, 103 (1948).
[11] R. Debot, Bull. Soc. Roy. Sci. Liège 12, 647 (1943).

depth of the subsurface latent image, reached by 2-minute treatment with 0.2 percent thiosulfate, to be only one lattice plane, and the depth of the internal latent image, reached by 2-minute treatment with 2 percent thiosulfate solution, to be not more than 10 to 15 lattice planes below the original surface.

III. HERSCHEL EFFECT AND LATENT-IMAGE STABILITY

Urbach and Wolinski[12] had found that changes in the latent image which occurred during the exposure or immediately after it could be expressed by a decrease of the Herschel effect, even if not detectable by other means; these changes were stated to extend up to 22 hours after the white-light exposure. It was found, however, in this investigation that a decrease of the Herschel effect at room temperature could be observed over long periods of time and also over a considerable length of time at higher temperatures. A decrease of the Herschel effect with keeping time was also mentioned in a review on the Herschel effect by Hautot and Falla.[8] In the present investigation, it was found that after 11 days at room temperature the Herschel effect decreased considerably. This can be seen from a comparison of Fig. 3 with Fig. 1, and the comparison also shows a simultaneous decrease of density for the higher exposures. There was not much difference between the keeping times of 11 days and 25 days, but Fig. 4 demonstrates the complete disappearance of the Herschel effect after 3 months, except for a remnant at high exposure, and a very strong density decrease in the high exposure region, resulting in a complete change of shape of the density curve.

It seems noteworthy that the decrease of the Herschel effect was accompanied in this case by a latent-image decay. The decrease of the Herschel effect can be interpreted as a growth of the development centers, so that a loss of one or two silver atoms would still leave them above the critical size necessary for development. This is substantiated by the findings of Burton and Berg,[2] who concluded that a tendency for growth exists for specks in the size range of the development centers. It seems hard to reconcile this phenomenon with latent-image decay. However, the two phenomena could be separated by raising the temperature during keeping time. After a 24-hour period of incubation at 110°F and 20 percent relative humidity, the Herschel effect was strongly decreased without any appreciable change in the density curve. After 3 days' incubation at 110°F and 20 percent relative humidity, the Herschel effect was still further reduced. Extension of the incubation time to 6 and 9 days did not produce any further change in the very slight Herschel effect, nor did 9 days' incubation at the higher temperature affect the shape or position of the density curve to any appreciable extent.

It seems that the growth of development centers at the surface is favored by the higher temperature, which could be expected; however, the density decrease at higher exposures is not obtained after 9 days' incubation at 110°F, not even to the extent observed after 11 days' keeping at room temperature. This does not seem in accord with an ex-

TABLE III. Development by D-16 and thiosulfate at 68°F.

% Thiosulfate in developer	Time of development	% Remaining AgBr	Depth of penetration, Δr
0.5	3 min.	81	26 mμ
1.0	6 min.	42	74 mμ
1.0	10 min.	26	108 mμ
1.0	14 min.	14	143 mμ
1.5	10 min.	9	160 mμ
1.5	14 min.	0.4	253 mμ

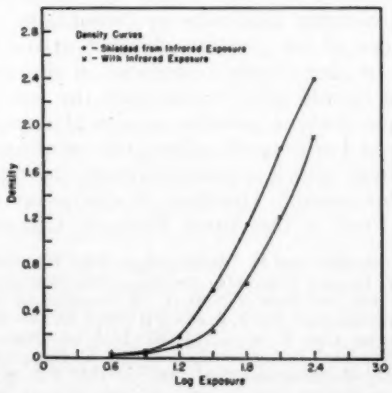

FIG. 5. Density curves with high-intensity white exposure. Development in D-16—4 min. at 68°F.

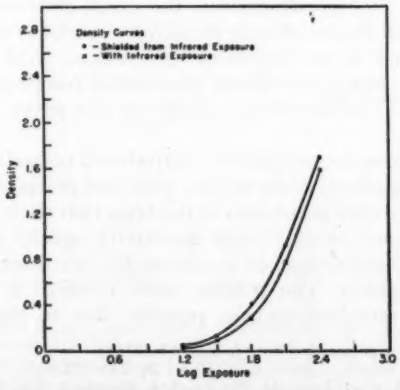

FIG. 6. Density curves with low-intensity white exposure. Development in D-16—4 min. at 68°F.

[12] F. Urbach and A. Wolinski, Wien. Akad. Ber. (IIa) **147**, 31 (1938).

BOSTON UNIVERSITY LIBRARIES

planation of the density decrease at higher exposure steps as a breaking-up of surface centers in favor of an increase in the internal centers. Any change of distribution based on ion migration should proceed more easily at higher temperatures. An explanation of the density decrease at higher exposure steps is lacking, therefore.

On the other hand, the information gained through the Herschel effect throws some light on the formation and stability of the development centers. Lescynski[13] stated more than twenty years ago that the same density produced at a low intensity level gave a much smaller Herschel effect than if produced by a high-intensity exposure. Recently, these findings were confirmed by Meĭklyar.[14] Burton and Berg plausibly interpreted this different behavior as due to the large number of small centers produced by high-intensity exposures. Figures 5 and 6 show very strikingly this different response, the very strong Herschel effect with high-intensity exposure (0.01 second) and the small effect with low-intensity exposure (342 seconds). In both cases, the Herschel effect decreased with 24-hour incubation at 110°F and 20 percent relative humidity by about 50 percent.

[13] W. F. Lescynski, Zeits. f. wiss. Phot. **24**, 275 (1926).

[14] P. B. Meĭklyar, Comptes Rendus U.R.S.S. **59**, 705 (1946).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

Some Remarks on Gold Treatment of Photographic Silver Halide Emulsions

F. W. H. MUELLER
Ansco Research Laboratories, Binghamton, New York
(Received March 25, 1949)

Earlier attempts of various investigators to increase the sensitivity of photographic emulsions by addition of colloidal silver, silver sulfide, or gold are reviewed. Koslowsky's discovery of the effect of certain gold complex salts as powerful sensitizers is described and a possible mechanism of the effect discussed. The importance of this effect for the manufacture of modern high speed and fine grain film as well as special x-ray emulsions is indicated.

DEMANDS for higher sensitivity of photographic emulsions and films preferably with finer grain to suit the various applications have continuously stimulated efforts of academic and industrial research laboratories to study the mechanism of silver halide activation and latent image formation.

Rather early evidence was obtained that photographic sensitivity is increased by the presence of ultra-microscopic particles on the silver halide grain. The question as to the exact chemical and physical nature of such sensitivity specks or centers remained in the realm of speculation, until Sheppard's[1] work established that sulfur compounds in gelatin produce silver sulfide on the silver halide grains.

Various investigators[2–4] introduced colloidal particles of silver, silver sulfide, gold, and platinum into photographic emulsions in the hope that such nuclei would act as additional sensitivity specks or enlarge ripening specks produced by the interaction with gelatin. The results, while positive in some cases, remained erratic, possibly due to the irreproducible properties of gelatin itself. For this reason, no commercial value could be attached to these otherwise interesting experiments.

Evidence, however, is accumulating that gold salts (not colloidal dispersions of the element) are being used in modern emulsion manufacture.[5–7] Yet it may not have been generally realized from some of these disclosures that it was the use of certain complex gold salts as "sensitizers" which led to one of the greatest advances in the manufacture of photographic emulsions in this country about a decade ago. At this time the use of this technique made it possible to raise film speeds by a factor of 4 or to produce fine grain emulsions with a sensitivity which seemed previously incompatible with this speed.* Therefore, a discussion of the earlier work in the Ansco Research Laboratories

[1] S. E. Sheppard, Phot. J. **65**, 380 (1925).
[2] W. Jenisch, Zeits. f. wiss. Phot. **24**, 248 (1926).
[3] B. H. Carrol and D. Hubbard, J. Research Nat. Bur. of Stand. **1**, 565–88 (1928).
[4] H. Weiss, Zeits. f. Physik. Chemie. **54**, 3051 (1906).

[5] R. Koslowsky and H. Mueller, Agfa Film Plant, Wolfen, Germany. Reports September–October 1936. *Bibliography of Scientific and Industrial Reports* (U. S. Department of Commerce, Washington), Vol. **8**, p. 873. PB 70053, fr. 831–50.
[6] H. Keller, Ger. P. Appl. 114,185 (2/17/43). *Bibliography of Scientific and Industrial Reports* (U. S. Department of Commerce, Washington), Vol. **10**, p. 158. PB 84,987 fr. 8366–7.
[7] Waller, Collins, and Dodd, United States Patent 2,399,083 (4/23/46).
* Ansco Superpan Supreme and Ansco Ultraspeed Pan Class 1-A Award of the Academy of Motion Picture Arts and Sciences in 1937.

might be in order. Furthermore, an elucidation of the mechanism of the gold salt treatment may lead to a better understanding of the nature of sensitivity specks.

In 1936 Koslowsky[5] of the Agfa Laboratories first demonstrated that a permanent two- to three-fold increase of sensitivity in gelatin silver halide emulsions could be achieved by incorporation of gold complex salts in the form of aqueous solutions of ammonium or alkali aurous thiocyanates during the second digestion or after ripening, or by adding such solutions to the melted emulsion before coating on a support (Figs. 1 and 2). Soon thereafter the writer found that a bathing treatment was similarly effective for both unexposed and exposed films (not published) (Fig. 3). Koslowsky assumed that the silver halide dissolving action of excess alkali thiocyanate and thiosulfate made available silver nuclei (sub-surface?) which were formed during the ripening process and which were then partially replaced by gold. This seemed to be in line with the observation that strong latensification by post-fixation gold treatment and physical development was obtained, a fact which was more recently confirmed by James.[8]

It is still believed that the sensitizing effect of gold salts upon silver halide emulsions—similar to that of mercury[9]—is more closely related to the presence of metallic silver nuclei or photo silver on the ripened emulsion grain rather than to the presence of silver sulfide. When gold salt sensitizers are added to an emulsion before digestion, only a negligible increase in speed is obtained presumably because of the absence of ripening silver.

Koslowsky noted that gold complex salts of very similar chemical structure, for instance, potassium aurous thiocyanate and potassium aurous cyanate behave photographically entirely different, the latter producing no sensitizing effect at all. During further investigation at the Ansco Laboratories, it became evident that complex salts, the complex ion of which has a dissolving action on silver as well as on silver halide and silver sulfide, are unsuitable for producing speed increases of practical significance. It is believed that in this case the silver of the ripening speck is dissolved either partially or completely before the "plating" reaction (replacement of silver with gold) can occur or that the gold specks are dissolved in the excess of potassium cyanide, used to stabilize the complex salt. On the other hand, gold complex salt solutions which will not dissolve silver, such as aurous thiocyanate or aurous thiosulfate, produce the desired increase in the sensitivity of the emulsion when added during the second digestion or when applied

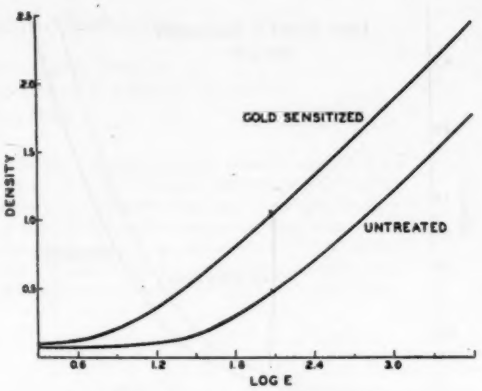

FIG. 1. Typical portrait emulsion (Ansco).

by bathing the finished film. This may serve as experimental evidence that silver must be present in sensitivity specks. The case of hypersensitization of unexposed film with gold salts lends further support to this conclusion.

The nature of the chemical reaction by which the gold plating occurs, is thought by the writer to be similar if not identical with that occurring in the old gold toning and fixing solutions[10,11] (which were used in connection with printing-out paper) expressed by the simple equation

$$Ag + Me' \longrightarrow Ag' + Me.$$
(noble metal ion)

The physical and chemical mechanism occurring on the grain can be thought to consist of a formation of a localized electrolytic cell. Erbacher[12] has shown that conditions for deposition of a noble

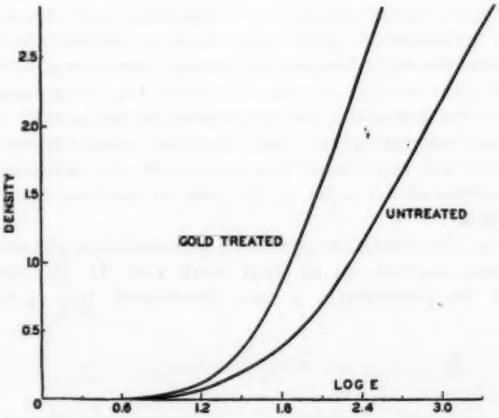

FIG. 2. Low speed, fine-grain emulsion (Ansco).

[8] T. H. James, *et al*, P.S.A. J. **14**, 349–53 (1948).
[9] F. Dersch and H. Dürr, S.M.P.E. **XXVIII**, 178–86 (1937).
[10] F. Formstecher, Phot. Ind. 874–5 (1921).
[11] J. M. Eder, *Handbuch d. Photographie* (1928), Vol. IV, I, pp. 42–63.
[12] O. Erbacher, Zeits. f. Electrochemie **38**, 534 (1932).

BOSTON UNIVERSITY LIBRARIES

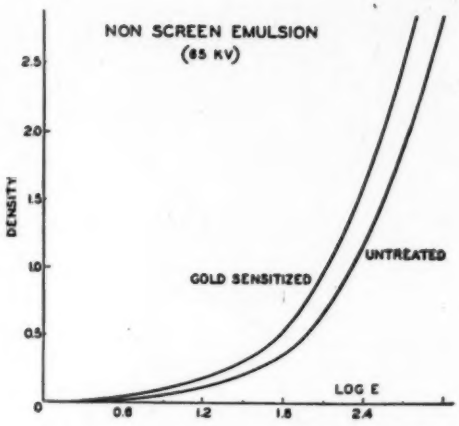

FIG. 3. Non-screen emulsion (65 kv) (Ansco).

salt on a less noble metal are particularly favorable for gold on silver and result in a multilayer plating. This is in good agreement with our photographic experience with other noble metal salt solutions of the platinum class which produced distinctly smaller sensitivity increases than those obtained with gold.

A number of experiments recently described by James[8] refer to latensification with gold salts and to the stability of gold-treated films towards certain bleaching agents, such as acid potassium dichromate; they are in excellent agreement with earlier work by the writer and, therefore, need not be described here again in detail. In other experiments at the Ansco Laboratories, it was possible to combine gold and mercury latensification or hypersensitization with an additive effect in certain emulsions without increasing the fog unduly. If amalgamation of silver specks enhances the electron-attracting properties of the sensitivity specks, a "gold-plated" speck may undergo similar amalgamation on subsequent mercury vapor treatment and may further increase the speck size or increase the electron-attracting properties of the speck. It is an interesting fact that the final result on such combined treatments was practically the same regardless of the order of the gold or mercury treatment.

In the course of some hypersensitizing experiments carried out in joint work with H. Hoerlin (to be published), it was discovered that gold sensitized emulsions designed for industrial radiography show a specifically increased sensitivity to high voltage radiation and thus record the impact of electrons in a manner superior to non-gold sensitized emulsions. It must be assumed that in this case too, the gold nuclei act as superior electron traps in the case of hypersensitization, or catalyze development in a more efficient manner in the case of latensification.

CONCLUSIONS

The existence of an electric potential in a complex gold salt solution in contact with metallic silver nuclei either in the form of ripening silver or photo silver is assumed to lead to a deposition of metallic gold *in situ*. Since this localized deposition of gold on the silver halide grain increases the emulsion speed substantially, gold nuclei must act as more efficient electron traps in terms of modern theories for the formation of sensitivity specks. That gold also serves as a more efficient catalyst in the development of exposed silver halide has been shown in the meantime by James.[13]

The relative ineffectiveness of colloidal gold dispersions as sensitizers may be explained by the lack of a potential. It seems possible that adsorption forces binding colloidal gold particles to the silver halide produce similar effects as the gold "plating" obtainable with ionic gold salts. There is no indication that colloidal gold has led to controllable speed increases of practical value. The formation of sensitivity specks, however, produced by deposition of gold from monovalent complex salts either during the ripening stage or when added as a coating final is a technique which has been utilized by manufacturers of photographic emulsions on a commercial scale to develop products of greatly improved sensitivity and finer grain.

ACKNOWLEDGMENT

As stated in the introduction, the author was but one of a number of contributors to the development of Koslowsky's discovery. The stimulating discussions and contributions of other members of the Ansco Laboratories, particularly Dr. H. H. Duerr and Dr. F. H. Dersch, are gratefully acknowledged.

[13] T. H. James, J. Coll. Science 3, 447 (1948).

Analysis of Para-Elliptic Reflector

BENJAMIN GOLDBERG AND NELS C. BENSON
Engineer Research and Development Laboratories, Fort Belvoir, Virginia
(Received April 1, 1949)

This paper presents an analysis of the para-elliptic reflector together with its possible applications when used as a searchlight mirror. The horizontal sections of a para-elliptic mirror are portions of ellipses and the vertical sections are parabolas. The mirror is so designed that each parabola lies in a vertical plane through the secondary focus of the horizontal ellipse which passes through the vertex of the mirror and each elliptical section lies in a plane parallel to that of the ellipse through the vertex of the mirror. A beam of light projected from such a reflector converges in the horizontal planes to a line common to the intersection of the vertical planes and then diverges as a spread beam horizontally with minimum spread vertically.

I. SUBJECT

THIS paper presents a mathematical analysis of the para-elliptic reflector together with its possible applications when used as a searchlight mirror.

II. INVESTIGATION

The horizontal sections of a para-elliptic mirror are portions of ellipses and the vertical sections are parabolas. When such a mirror is intended for use in a searchlight for projecting a beam of light, it is so designed that:

(a) Each parabola lies in a vertical plane through the secondary focus of the horizontal ellipse which passes through the vertex of the mirror.
(b) Each elliptical section lies in a plane parallel to that of the ellipse through the vertex of the mirror.
(c) The focus of the parabola through the vertex of the mirror corresponds with the primary focus of the ellipse through the vertex of the mirror.

Light emanating from the primary focus of a mirror which is a segment of an ellipsoid of revolution will be reflected as a cone of light which converges to a point at the secondary focus. From the secondary focus, the light diverges as a cone beam. These properties are illustrated in Fig. 1. In this case, the periphery of the mirror is a circle and the intersection of the mirror surface with any plane which passes through the primary focus and vertex of the mirror is an ellipse.

If the horizontal elements of the ellipsoid could be retained and the vertical elements bent back and made to conform to parabolas without changing the foci, then the rays of light which originally passed through the secondary focus would be unaltered in their azimuth angles but would be rotated upward and downward until parallel to the horizontal plane through the optical axis. Light striking all the elements of any one vertical parabola would be reflected in a vertical line passing through the secondary focus. The total effect from such a mirror would be a fan of vertical sheets intersecting in a vertical line passing through the secondary focus as illustrated in Fig. 2. This effect cannot be absolutely obtained since there is no standard equation representing this surface. The nearest approach to the above effect is produced by the so-called para-elliptic mirror. Light projected from a para-elliptic mirror diverges somewhat from the pattern shown in Fig. 2 because each vertical parabola has its own unique focal point, and a light source of finite size is used.

A para-elliptic reflector with primary focus at F, secondary focus at F_1, and optical axis on the X axis, together with a system of coordinates consisting of three mutually perpendicular planes is shown in Fig. 3.

The equation of the ellipse through the optical axis in the XY plane is:

$$x^2/a^2+y^2/b^2=1 \qquad (1)$$

an arc of this ellipse is represented by USRV. The lengths of the semimajor and semiminor axes "a" and "b" are dictated by the application in which the mirror is to be used and depend upon the selection of the position of the primary and secondary foci, together with the angle of convergence of the light rays required to satisfy the particular application.

The equation of the parabola through the optical axis in the YZ plane an arc of which is represented N_1SN is:

$$z^2=4(FS)(x+a). \qquad (2)$$

The focal length, FS, of the parabola is identical to

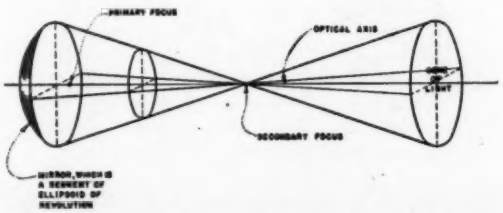

FIG. 1. Beam produced by mirror which is a segment of ellipsoid of revolution.

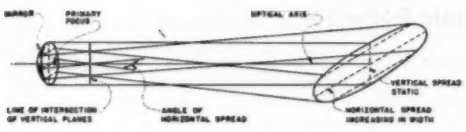

FIG. 2. Type of beam produced by a para-elliptic mirror.

the primary focal length of the ellipse through the optical axis.

By hypothesis, the periphery of the mirror, VN_1UN, is an ellipse with major axis NN_1 and the minor axis UV, and it lies in a plane parallel to the YZ plane. The equation of this ellipse, with axes Z_1 and Y_1 through the center of the ellipse, is:

$$z_1^2/(NA)^2+y_1^2/(UA)^2=1. \tag{3}$$

The equation of the variable parabola WRP with the origin at point R, for the variable point P, may be written:

$$(PR_1)^2=4f(RR_1)$$

where f represents the varying focal lengths of the parabolas as point R moves along the elliptical arc $(USRV)$ through the vertex of the mirror. The length PR_1 is the value z_1 of the equation of the periphery of the mirror. Therefore, the value of z_1 of the periphery of the mirror may be expressed as follows as point R, moves from A to V in the mirror periphery:

$$z_1^2=4f(RR_1). \tag{4}$$

RR_1 varies as point R moves along the elliptical arc $(USRV)$ through the vertex of the mirror. RR_1 can be evaluated as follows:

$$RR_1=F_1R-F_1R_1.$$

$F_1R=[(F_1T)^2+(TR)^2]^{\frac{1}{2}}$ since F_1R is the hypotenuse of the right triangle RTF_1

$$F_1T=2a-F_1D-ST;$$

but $F_1D=FS$ and $ST=OS-OT$. Since, for the variable point R, $OT=x$, $OS-OT$ is equal to $a-x$. Then

$$F_1T=2a-FS-(a-x)=a-FS+x.$$

Since both factors of the expression $a-FS$ are constants for any specific para-elliptic mirror, the expression $a-FS$ may be called the constant m. Then

$$F_1T=m+x$$

$TR=y$ of the point R on the elliptical arc $USRV$. Therefore, $F_1R=[(m+x)^2+y^2]^{\frac{1}{2}}$. $F_1R_1=[(F_1A)^2+(AR_1)^2]^{\frac{1}{2}}$, since F_1R_1 is the hypotenuse of the right triangle R_1AF.

$$F_1A=2a-F_1D-SA;$$

but $F_1D=FS$. Then $F_1A=2a-FS-SA$. Since all factors of the expression $2a-FS-SA$ are constants for any specific para-elliptic mirror, the expression $2a-FS-SA$ may be called the constant n. Then

$$F_1A=n$$

$AR_1=y_1$, the coordinate of point P on the mirror periphery. Therefore, $F_1R_1=(n^2+y_1^2)^{\frac{1}{2}}$ and

$$RR_1=[(m+x)^2+y^2]^{\frac{1}{2}}-(n^2+y_1^2)^{\frac{1}{2}}. \tag{5}$$

In order to solve the above equation for RR_1, it is necessary to establish a relationship between y and y_1. This relationship, from Fig. 3, is: Since corresponding angles of the two triangles RTF_1 and R_1AF_1 are equal:

$$y_1/y=F_1A/F_1T=n/m+x. \tag{6}$$

Substituting $[a^2-(a^2y^2/b^2)]^{\frac{1}{2}}$ for x (see Eq. (1)) and solving for y in terms of y_1:

$$y=\frac{2mn/y_1\pm[4m^2n^2/y_1^2-4(n^2/y_1^2+a^2/b^2)(m^2-a^2)]^{\frac{1}{2}}}{2(n^2/y_1^2+a^2/b^2)} \tag{7}$$

Substitution of various values of y_1 in the above expression will yield corresponding values of y which can be used in the RR_1 equation.

By use of the equations (1 through 7) set forth the changing focal length, f, for a particular para-elliptic mirror can be determined.

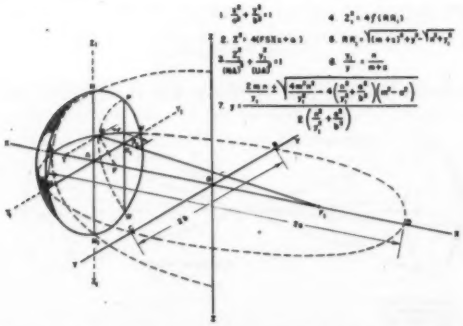

FIG. 3. Illustration of the elements of a para-elliptic reflector with a system of coordinates having origin at center of the ellipse through the optical axis.

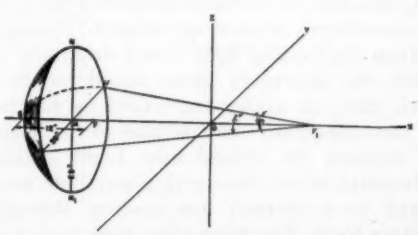

FIG. 4. Illustration of para-elliptic mirror with focal length of 12 inches, short diameter 20 inches, and angle of convergence of 17 degrees.

TABLE I. Evaluation of "f" values of y_1 from 1 through 9.

1	2	3	4	5	6
y_1	$\dfrac{\frac{2mn}{y_1}+\left[\frac{4m^2n^2}{y_1^2}-4\left(\frac{n^2}{y_1^2}+\frac{a^2}{b^2}\right)(m^2-a^2)\right]^{\frac{1}{2}}}{2\left(\frac{n^2}{y_1^2}+\frac{a^2}{b^2}\right)}=y$	$\dfrac{ny}{y_1}-m=x$ $66.911\dfrac{y}{y_1}-28.716=x$	RR_1	$\left(1-\dfrac{y_1^2}{(10)^2}\right)(11)^2=z_1^2$	$\dfrac{z_1^2}{4(RR_1)}=f$
1	1.0373	40.691	2.495	119.79	12.004
2	2.0722	40.610	2.416	116.16	12.019
3	3.1025	40.477	2.285	110.11	12.047
4	4.1256	40.296	2.105	101.64	12.071
5	5.1396	40.062	1.872	90.75	12.119
6	6.1422	39.781	1.593	77.44	12.153
7	7.1313	39.446	1.258	61.71	12.263
8	8.1050	39.062	0.874	43.56	12.460
9	9.0610	38.643	0.453	22.99	12.688

Consider a para-elliptic mirror having a vertical diameter of 22 inches, a horizontal diameter of 20 inches, a focal length of 12 inches, and an angle of convergence of 17° as illustrated in Fig. 4. The z value at the mirror periphery of the parabola through the optical axis (Eq. (2)) is 11 inches; and the NA and UA values for the equation of the mirror periphery (Eq. 3), are 11 inches and 10 inches, respectively.

From equation for the parabola through the optical axis, $z^2=4(FS)(x+a)$; $(x+a)=2.521$ inches for the point at the mirror periphery. This is also the value of SA, since, in Fig. 3, a is the distance OS, and the x value for the mirror periphery is OA and the difference between OS and OA is SA.

The values of a and b of the equation for the ellipse through the optical axis can now be determined.

$$\text{Tan. } 8\tfrac{1}{2}^\circ = 10/2a - FS - SA = 10/2a - 12 - 2.521$$
$$a = 40.716 \text{ inches};$$

but $x+a=2.521$ inches. Then $x=-38.818$ inches for the point at the mirror periphery. With substitution of the x, y, and a values in the equation of the ellipse through the axis; $x^2/a^2+y^2/b^2=1$, where $b=28.818$ inches, $m=a-FS=40.716-12=28.716$, and $n=2a-FS-SA=81.432-12-2.521=66.911$ inches.

The RR_1 value where $y_1=0$ is equal to SA or 2.521 inches. The RR_1 value where $y_1=10$ is 0. To determine the intermediate values, $y_1=1$ inch through $y_1=9$ inches, the various equations above are used in the following manner:

The particular value of y_1 in question is substituted in Eq. (7), which yields a corresponding value for y. The y_1 and the y values are substituted in Eq. (6), and the value of $m+x$ is obtained. RR_1 of Eq. (5), is determined by use of the values for y_1, y, and $m+x$.

The z_1 value of Eq. (4), corresponding to the y_1 in question, is found from Eq. (3). With the RR_1 and z_1 values of Eq. (4), now determined, that equation can be solved for "f" (the focal lengths of the variable parabolas).

Solutions were obtained for RR_1 and "f" for values of y_1 from 1 through 9 (see Table I). The plot of the RR_1 values against the y_1 values for this particular para-elliptic mirror is shown in Fig. 5. The plot of the "f" values against the y_1 values is shown in Fig. 6. This latter plot reveals the increasing focal lengths of the variable parabolas for the specific para-elliptic mirror resulting from the assumed dimensions and angle of convergence.

In order to determine the behavior of the reflected beam in the vertical planes, it is necessary to know the relation of the focal length of each parabola to the distance from each parabola to the fixed light source placed at point F (Fig. 4), the focus of the para-elliptic mirror. If the distance

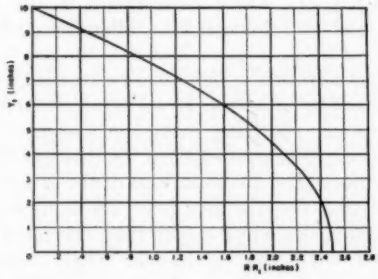

FIG. 5. Plot of the RR_1 values against the distance from the center of the specific para-elliptic reflector resulting from the assumed dimensions and angle of convergence.

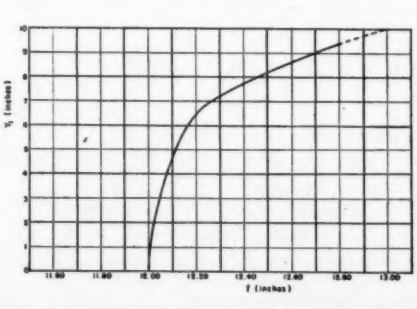

FIG. 6. Variation of the focal lengths of the parabolas in passing from the center to the edge of the specific para-elliptic reflector resulting from the assumed dimensions and angle of convergence.

BOSTON UNIVERSITY LIBRARIES

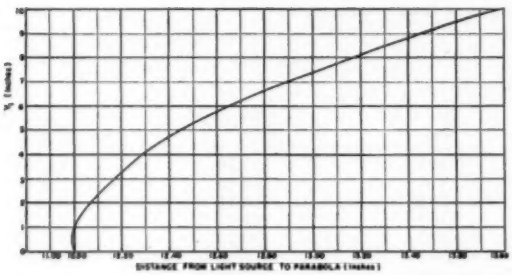

FIG. 7. Variation of the distance from the light source to the parabolas in passing from the center to the edge of the specific para-elliptic reflector resulting from the assumed dimensions and angle of convergence.

from the light source to any particular parabola is shorter than its focal length, the reflected beam will diverge in the plane of that parabola, and if the distance from the light source is longer than the focal length of a particular parabola, the beam will converge in the plane of that parabola. The distance from any parabola from the focus at F may be found as follows:

From Fig. 3, the distance $FR=[(RT)^2+(FT)^2]^{\frac{1}{2}}$ since RTF is a right triangle. $RT=y$, corresponding to the particular y_1 value which may be assumed for Eq. (7).

From Fig. 3, $FT=OT-OF$, but OT is the x value of Eq. (1), which corresponds to the y value above, and $OF=-a+FS$. Therefore, $FR=[y^2+(x+a-FS)^2]^{\frac{1}{2}}$. Figure 7 shows the variation of the distance from the light source to the parabolas for values of y_1 from 0 to 10 for the particular para-elliptic mirror resulting from the assumed dimensions and angle of convergence. In this specific case, it will be noted, by a comparison of the graph shown in Fig. 6 that the distance from the light source to any parabola is always longer than the focal length associated with that parabola. This indicates a somewhat convergent beam in the planes of the parabolas. The maximum convergence will be approximately three degrees.

III. DISCUSSION

The para-elliptic mirror is particularly useful, where it is desirable to protect the searchlight behind a barricade, for transmitting the light through a slit at the secondary focus of the ellipse. After passing through the slit, the light diverges into a horizontally fan-shaped beam, illuminating a wide area without wasting light skyward. The angle of spread will be determined by the relative positions of the foci of the ellipse. In the particular mirror cited in this paper, the slit would be almost 69½ inches from the mirror in the searchlight. In order to accommodate all the light from the mirror as it passes through the secondary focus, the slit should be about 2½ inches wide and 24 inches long.

If a searchlight, containing a para-elliptic mirror, were used from a casemate or tank, the slit—the only exposed portion—would make a small target.

To manufacture a para-elliptic mirror to meet the limitations of a particular application and to provide the desired beam distribution, it is necessary to calculate the physical dimensions required. For example, a tank installation may impose restrictions as to the size of mirror, the distance from the mirror to the aperture in the tank, and the size of the aperture; whereas another application may require a larger mirror placed at a greater distance from the aperture through which the light is emitted. With such limitations determined, the equations set forth in this paper provide a means of calculating all dimensions essential for the manufacture of the particular para-elliptic reflector desired and indicating how a cutting tool would have to be varied as it cuts the varying vertical parabolas on the reflector surface.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 39, NUMBER 6 JUNE, 1949

Emulsion Calibration Scale for Quantitative Spectroscopic Analysis

HAROLD K. HUGHES AND R. W. MURPHY
Socony-Vacuum Laboratories, Technical Service Department, Brooklyn 22, New York
(Received April 5, 1949)

The time devoted to collecting and plotting emulsion calibration data can be reduced considerably by the use of a new transmittance scale which is logarithmic from 1 to 30 percent and which then expands non-uniformly to 97 percent. Data are presented from seven independent laboratories which establish the approximate linearity of the emulsion calibration curves for Eastman SA-1 films, SA-1 plates, SA-2 films, I-N plates, and 33 plates. The practical uses of this scale on plotting boards and in a special emulsion calibration paper are discussed.

In certain cases, all emulsion calibration procedures can be eliminated. A transmittance slide rule for use in the quantitative determination of tetraethyl lead in gasoline is described.

I. INTRODUCTION

ONE of the important steps in quantitative spectroscopic analysis with other than photoelectric instruments is the calibration of the photographic emulsion at one or more wave-lengths adjacent to the analytical and standard lines. While several methods are well known, the whole process of collecting and plotting the required data by any of them is an appreciable fraction of the total time devoted to analysis. A simplification of emulsion calibration and even its entire elimination is desirable from the standpoint of reducing the cost of analyses and of speeding the issuing of reports.

This paper describes improved methods of handling the calibration process which reduce the number of data required, extend the useful exposure range, and provide a ready check on the accuracy of photometric readings. In certain cases, the calibration process may be eliminated completely. With the use of a new transmittance scale, emulsion calibration curves are essentially linear over the range of transmittance from 2 to 97 percent.

The emulsion calibration curve which is usually plotted in spectroscopy laboratories has as ordinate the logarithm of the transmittance (or some related function, such as the absorbance) and as abscissa the logarithm of the exposure (or some function proportional to it). With these scales, the response of photographic emulsions is the familiar sigmoid curve first determined by Hurter and Driffield and frequently named for them. It is approximately linear over the range of transmittance from 2 to 30 percent.

The ratio of the ordinate to abscissa scaling factors is often chosen so as to produce an emulsion calibration curve whose linear portion has a slope of 45 degrees for a gamma of two. This is a good compromise for the common emulsions employed in applied spectroscopy.

The linear portion of a calibration curve does not cover the range of most accurate photometric data, and thus, it is generally necessary to work also on the curved portions. Some laboratories report useful analyses with transmittances as low as 2 percent, while backgrounds in the vicinity of 97 percent are common. At transmittance values exceeding 30 percent, the curve flattens out and the accuracy with which exposure ratios can be obtained is reduced. This effect is so marked that Harvey[1] has proposed the use of a $\log(1-\tau)$ scale above $\tau=50$ percent. The slope of the $\log(1-\tau)$ *vs.* $\log\Sigma$-curve increases with increasing Σ. The Greek letters τ and Σ represent the transmittance and exposure, respectively.

Solarization in common spectroscopic emulsions is slight, even at $\tau=2$ percent, and curvature in this region is generally negligible.

II. LINEAR CALIBRATION CURVES

It is possible to produce an emulsion calibration curve which is essentially linear over the entire range of τ by modifying the ordinate scale, leaving the abscissa scale proportional to $\log\Sigma$. The requirements on the modified scale are that it must be analytically continuous, between 2 percent and approximately 30 percent τ it must be logarithmic, and above 30 percent it must expand, in comparison with a logarithmic scale, at such a rate that the resulting calibration curve is a straight line.

One way to approach this problem would be

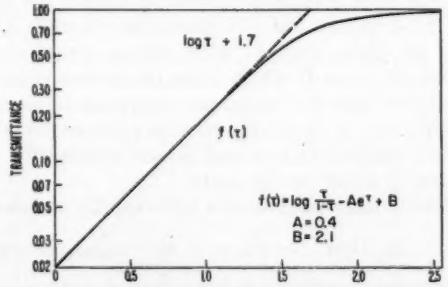

FIG. 1. Comparison of $\log\tau$- and modified τ-scales.

[1] C. E. Harvey, *A method of semi-quantitative spectrographic analysis* (Applied Research Laboratories, Glendale, California, 1947).

TABLE I. Table of τ-coordinate distances.

τ(%)	(Inches down) Board[a]	Paper[b]
2	0.00	0.00
5	1.88	1.60
10	3.30	2.81
20	4.74	4.03
30	5.59	4.76
40	6.23	5.30
50	6.77	5.76
60	7.26	6.18
70	7.80	6.64
80	8.51	7.24
82	8.69	7.39
84	8.87	7.55
86	9.12	7.76
88	9.40	8.00
90	9.72	8.27
91	9.92	8.44
92	10.13	8.62
93	10.38	8.83
94	10.67	9.08
95	11.01	9.37
96	11.45	9.74
97	12.00	10.21

[a] Calculated to fit the vertical scale of the ARL-calculating board.
[b] Calculated to fit standard 11×16-in. paper.

through a study of the numerous equations which have been devised to represent the plot of τ *versus* Σ. A review of these may be found in Mees.[2] While attractive, this approach was abandoned early in our investigation in favor of the method of curve-fitting an analytic expression to a number of experimental emulsion calibrations. Of a variety of functions investigated, the following has proved to be most satisfactory,

$$f(\tau)=\log\tau/(1-\tau)-0.4e^{\tau}. \qquad (1)$$

This function, which is shown in Fig. 1, has only one arbitrary constant, though in applying it to the construction of the vertical scale of a plotting board or in computing the positions of the ordinate lines on the special coordinate paper described below, it is necessary to add a positioning constant and to multiply by a scaling factor.[3]

The use of this new scale is illustrated in Fig. 2, where several emulsion calibration curves for Eastman SA-1 plates and I-N plates are plotted. All of these are quite straight with the exception of the portion of curve *D* which joins the lowest point. It is believed that this apparent exception arises from the difficulty of obtaining reliable photometric data at high transmittances and is not necessarily evidence of a defect in the scale.

Several other laboratories have kindly assisted in tests of this new scale. Their data are shown in Figs. 3a and 3b. Not shown in these 3 figures are data obtained in the authors' laboratory on a IV-N plate at 6141A and a Process plate at 2544A. In both cases, the approximate linearity of the resulting curves is clearly established.

From this test of the modified τ-scale on seven common photographic materials used in applied spectroscopy, covering a gamma-range of 0.9 to over 3, and a wave-length range from 2343A to 6141A, it is concluded that in most cases it is valid to draw a straight calibration curve. Whether the slight remaining curvature in a few instances is to be attributed to the data or to the τ-scale is still to be determined. This slight curvature never flattens any portion of the curve to such an extent that a $(1-\tau)$ curve is necessary.

III. THE CONVERSION OF PLOTTING BOARDS

Plotting boards which now employ a logarithmic ordinate scale may readily be converted to the new τ-function scale by pasting over the present one a piece of heavy paper or cardboard upon which ruled lines are spaced, as given in Table I. The second column has been computed particularly for the board manufactured by Applied Research Laboratories, which is the one in use in the authors' laboratory. For other sized boards, multiply this column by the appropriate scaling factor. Intermediate lines are located with the aid of a normal logarithmic scale aligned at two succeeding computed lines.

IV. EMULSION CALIBRATION PAPER

Since plotting boards are rather cumbersome, the authors have found it convenient to prepare a special emulsion calibration coordinate paper using the new τ-function. The ordinate runs from 2 to 97 percent and the abscissa from 1 to 25.6. The latter number is equal to $(1.5)^8$ and is for use with a

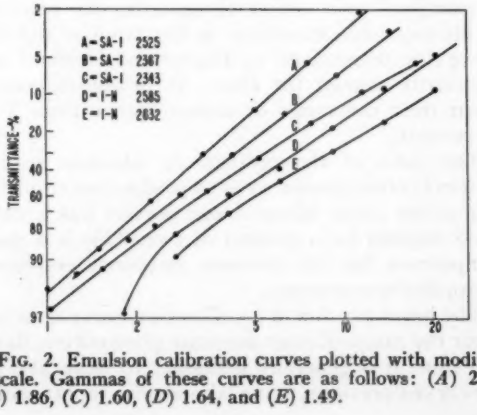

FIG. 2. Emulsion calibration curves plotted with modified τ-scale. Gammas of these curves are as follows: (*A*) 2.15, (*B*) 1.86, (*C*) 1.60, (*D*) 1.64, and (*E*) 1.49.

[2] C. E. K. Mees, *The Theory of the Photographic Process* (The Macmillan Company, New York, 1942).
[3] Following the publication of the abstract of this paper in the program of the winter meeting of the Society, March 10–12, 1949, C. Feldman called the authors' attention to the work of H. Kaiser, Spectrochem. Acta 3, 159 (1948). This reference has not been available, but it is understood that Kaiser uses the first term on the right of Eq. (1).

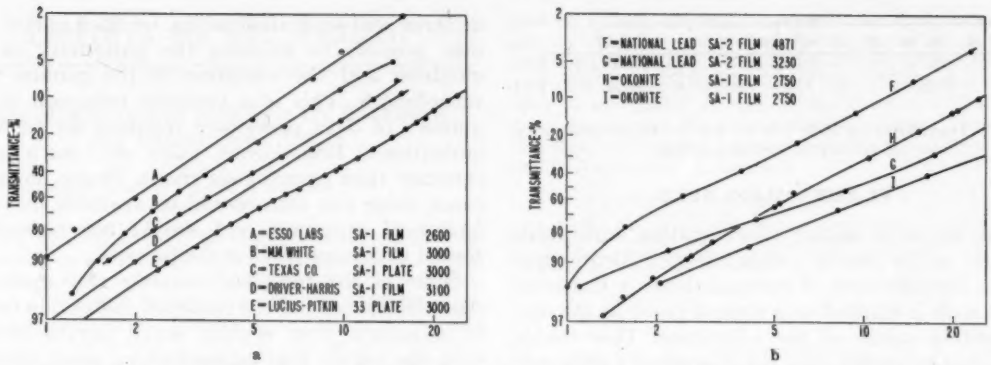

FIG. 3. Emulsion calibration curves plotted with modified τ-scale. Gammas of these curves are as follows: (*A*) 1.68, (*B*) 1.59, (*C*) 1.48, (*D*) 1.57, (*E*) 1.36, (*F*) 1.24, (*G*) 0.91, (*H*) 1.28, and (*I*) 0.96.

step-sector of ratio 1.5. The original drawing of the coordinate paper was reproduced by the photo-offset process. The third column of Table I gives the distances for use on 11×16-in. paper. The vertical lines were traced from standard 2-cycle paper, 10 in. to the cycle. The binding margin is 1 in.

The ratio of the scaling factors is 2.47. Gammas are computed by multiplying this number by the slope. A sample of this paper will be supplied on request.

Preliminary curves are best plotted on a separate sheet of log-log paper. However, it is possible to use emulsion calibration paper by doubling the ordinates from 2 to 50 percent. It is convenient to write these values on the right-hand margin. This gives a scale from 4 to 100 percent which is used with one cycle of the abscissa.

These details relate to the employment of the new τ-scale that is considered most convenient by the authors. Other laboratories may find different ranges and scaling factors more suitable.

V. ELIMINATION OF THE EMULSION CALIBRATION

The reduction of the emulsion calibration curve to a linear relation between the τ-function and the log exposure permits the elimination of the specific determination of the emulsion calibration. How this is done may be seen by referring to Fig. 4 where an emulsion calibration of slope γ is drawn.

The logarithm of the ratio of unknown and standard line exposures, Σ_u and Σ_s, is

$$\log(\Sigma_u/\Sigma_s)=1/\gamma\{f(\tau_u)-f(\tau_s)\}. \qquad (2)$$

If, instead of the correct emulsion calibration, any arbitrary straight line of slope γ' is used, a new value is obtained,

$$\log'(\Sigma_u/\Sigma_s)=1/\gamma'\{f(\tau_u)-f(\tau_s)\}, \qquad (3)$$

which differs from the true log ratio by the factor γ/γ'.

The new values, $\log'(\Sigma_u/\Sigma_s)$, can be plotted against log concentrations of standard samples, giving an analytical curve without the use of an emulsion calibration. In employing this analytical curve it is only necessary to insure that the true γ of the emulsion does not change, and that the same arbitrary line of slope γ' is used for computing $\log'(\Sigma_u/\Sigma_s)$.

It is convenient for computation to prepare a sliding concentration scale from the analytical curve. This scale replaces the log exposure scale of the τ-function coordinate paper. On it is marked the concentration of an element in terms of the difference between the abscissas corresponding to the element and internal standard line transmittances.

When an internal standard is unnecessary in a determination, the concentration of an element may be substituted for the log exposure scale. Standard samples establish a curve from which the concentration may be read directly in terms of transmittance.

If the analytical curve of exposure *vs.* concentration is a straight line on log-log paper, then it will also be straight when plotted as transmittance *vs.* concentration on emulsion calibration paper.

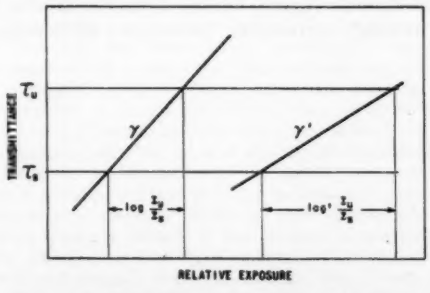

FIG. 4. Substitution of an arbitrary line for the emulsion calibration.

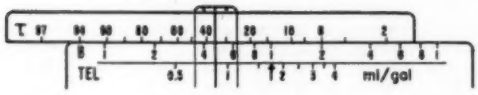

FIG. 5. Transmittance slide rule for use in the determination of tetraethyl lead in gasoline.

VI. THE τ-SLIDE RULE

The use of a sliding concentration scale leads directly to the idea of a slide rule as a further step in the simplification of computations. A transmittance scale is marked on a strip of paper at the corresponding values of the τ-function. This τ-scale may then be pasted over the A-scale of a slide rule as shown in Fig. 5.

Another scale pasted on the slide is laid off in terms of concentration of an element. To determine an element concentration, the index of this scale, marked with an arrow in Fig. 5, is set at the transmittance of the internal standard. The concentration is then read off directly below the corresponding unknown element transmittance.

The B-scale of the slide rule is convenient to use in preparing the concentration scale. Setting the center index of B at the internal standard transmittance, values of B for standard samples of the unknown are read at the transmittances of the unknown element. By graphical interpolation, a scale on the slide is marked off in concentration adjacent to the appropriate B values.

VII. DISCUSSION

The difficulty of obtaining reliable calibration data is well known. No review will be made here of the effect of non-uniform irradiation of the slit, the intermittency effect, the errors introduced by background and incorrect "100-setting" of the photometer, the Schwarzschild-Villiger effect, and other aberrations. These all contribute to the small uncertainties in the evaluation of the new scale. It is worth noting, however, that the τ-scale now provides a ready test of photometric readings, once the linearity of the emulsion calibration curve is established. This is particularly valuable in the high transmittance region where data heretofore have been notably unreliable. Moreover, with only two or three well-separated points, for each curve, it is now possible to examine the uniformity of the emulsion and the variation of the gamma with wave-length. This is a valuable reduction in the number of data previously required for such examinations. Homologous pairs are much more common than homologous triads. Hence, in many cases, these test data should be available from the analytical exposures and should not necessitate special exposures just for the purpose.

Some of the foregoing remarks rely upon the essential linearity of the emulsion calibration curve. It is known from several dozen curves plotted with the τ-scale that occasionally a small residual curvature is left which may be truly representative of the emulsion and not ascribable to poor data. The extent to which this departure from a straight line can be tolerated depends, of course, upon the analytical precision sought and other considerations. Normally, one would expect larger errors in the final analytical results from other causes, such as variations in the source and in the emulsion.

The authors believe that the use of this new τ-scale and the methods of eliminating the calibration process described here have wide applicability in applied spectroscopy and that they will thus simplify quantitative analyses.

VIII. ACKNOWLEDGMENTS

This work has been done under the direction of J. B. Rather, Jr., Assistant Director of the Laboratory, and of R. W. Moore, Supervisor of the New Methods, Physics, and Metallurgical Section. It is a pleasure to acknowledge their encouragement. Our colleague, J. W. Anderson, assisted in collecting some of the data.

We are grateful to several other laboratories for providing data with which to check the new scale, and, in particular, wish to thank the following: C. L. Guettel of the Driver-Harris Company; E. L. Baldeschwieler and J. Hansen of the Standard Oil Development Company; R. H. Bell and C. A. Jedlicka of Lucius-Pitkin, Inc.; E. J. Dunn, Jr., and A. J. Mittledorf of the National Lead Company; M. E. Lampman of the Okonite Company; H. Levin and R. Pomatti of The Texas Company.

Letters to the Editor

Standard Response Functions for Protanopic and Deuteranopic Vision

DEANE B. JUDD
National Bureau of Standards, Washington 25, D. C.
March 28, 1949

IT has been suggested by Le Grand[1] that the standard response functions, proposed by me[2] to represent protanopic and deuteranopic vision, introduce contradictions in the calculation of luminous quantities because two of the three primaries are unassociated with luminosity. He points out that if we explain protanopia by assuming that the primary carrying all the luminosity is suppressed, this would seem to require the observer to be not merely red-green blind, but completely blind. He recommends that another coordinate system, equally or perhaps better in accord with the color confusions of deuteranopes and protanopes, whose W_p process has a luminosity coefficient of about 13 percent, be used "in order to avoid contradictions."

Leaving out of account possible deviations from the law of additivity of luminances on which photometry is based (a law by no means certainly established),[3] a consistent photometric scale may be derived by computation based on the luminosity function. If the luminosity function is correct for the observer, he will make photometric matches that do not contradict the computed values. Since the standard luminosity function, and those proposed by me for protanopic and deuteranopic vision alike, all fail precisely to represent any actual observer, the use of the scales resting on these functions all will disclose contradictions, some not at all negligible. The validity of the standard scale of luminance is simply of an arbitrary sort based on definition. The usefulness of the definition depends upon the fact that the standard luminosity function is intermediate to those of observers having normal vision. The usefulness of the luminosity functions proposed by me to represent dichromatic vision, and that suggested by Le Grand alike, also arises from the fact that they all lie intermediate to the luminosity functions of observers having the corresponding type of vision. No contradictions will be avoided by using the less convenient function proposed by Le Grand.

It might seem from the foregoing that there is no point to the note written by Le Grand; but that is not true either. The question of how the luminosity process in protanopic eyes manages to be transferred from receptor cells responding in accord with the normal luminosity function, to those responding in accord with the protanopic luminosity function, is a difficulty in visual theories assuming separate receptors for the luminous and chromatic aspects of the response (the so-called dominator-modulator type of theory suggested by the researches of Granit).[4] In addition to the suppression of the process yielding the normal luminosity function, it is necessary to assume that the W_p process takes over the role of dominator, in order to account for protanopic vision. This is an *ad hoc* assumption detracting from the plausibility of this particular form of dominator-modulator theory.

If, however, we take instead of the normal luminosity function, W_d, to apply to deuteranopic vision, the function, W_d', preferred by Le Grand, no such *ad hoc* assumption is required. The suppression of the W_d' function leaves the system with 13 percent of its normal luminosity distributed correctly according to the protanopic luminosity functions determined experimentally. This does, indeed, remove the difficulty spoken of by Le Grand as a "contradiction."

The difference between the two luminosity functions, W_d and W_d' is much smaller than the spread of luminosity functions for normal observers. My view is that the question will only be settled by measuring the luminosity functions of hundreds of normal and deuteranopic observers under identical conditions, and the application of careful statistical treatment to the results. It may be that there is a significant difference between the normal and deuteranopic luminosity functions, but until this is proven the computationally simpler choice, W_d, may as well be used.

[1] Yves Le Grand, J. Opt. Soc. Am. **38**, 815 (1948).
[2] D. B. Judd, J. Opt. Soc. Am. **35**, 199 (1945).
[3] A. Dresler, Das Licht **7**, 203 (1937); A. Kohlrausch, Das Licht **5**, 259, 275 (1935); J. Urbanek and E. Ferencz, International Beleuchtungskommission, Proc. 10th Session, Scheveningen 1939 **1**, 44 (1942).
[4] R. Granit, Nature **151**, 11 (1943).

Non-Intermittent Motion Picture Projection

WILLIAM C. PLANK
1213 Maple Avenue, Vallejo, California
March 9, 1949

THE numbers and the needs of amateur motion picture projectionists make expedient a restatement or recapitulation of the advantages of non-intermittent motion picture projectors. This is thought necessary on account of the want of due appreciation of the useful characteristics of this type of projector, which the literature published on the subject has revealed.

In a paper entitled "The Problem of Motion Picture Projection from Continuously Moving Film," which was reprinted in the January 1933 number of the J. Soc. Mot. Pict. Eng., a summary of the advantages of non-intermittent projectors was published, which is notable for the number and the importance of those omitted.

Another paper which is believed to undervalue the utility of non-intermittent projectors bears the title, "Is the Continuous Projector Commercially Practical?" (Transactions No. 18 Soc. Mot. Pict. Eng.)

The purpose of this letter is to submit, for the notice of those who require a more adequate amount of data for an appraisal, several additional advantages which were previously passed unnoticed.

In non-intermittent projectors an undisputed advantage is that instead of curtailing the exposure, or stationary period of the image, an additional length of time is assigned to it. The depth, or third dimensional effect brought out by non-intermittent projectors, is believed to be due to the advantage of the lap-dissolve.

Of the useful properties inherent in non-intermittent projectors, one of the most curious is the movement exhibited during a slow lap-dissolve of two successive picture images. As only two film frames take part in the movement, the apparent intermediate positions do not exist in the photographic record; but they are, nevertheless, conveniently introduced by the magic of the lap-dissolve. These movements can sometimes be stopped and started again more than once. In close-ups, the turning of the head or body is often particularly well rendered, with a life-like uniform motion no matter how slowly the change be made. The cause of this curious phenomenon is to be found in the irradiation of the retina. Its greatest advantage to amateurs is that it forms the basis for a successful and inexpensive means of exhibiting slow motion, when the rate of projection is reduced to two or three frames a second, and the taking rates have been at twenty-four or over. The unique slowed motions displayed, and the theory upon which they are based, provide a distinct source of interest and amusement to amateurs.

BOSTON UNIVERSITY LIBRARIES

This useful property of the lap-dissolve is also an important factor in the successful employment of lower rates of projection for amateurs. Many amateur shots can be projected successfully by the non-intermittent projector at rates of from twelve to eight frames a second.

A major advantage, involving higher orders of precision in motion picture registration, is that the registration of the non-intermittent projector can be made as independent of the inaccuracies of the film perforations as the sound track.

The effect of vibration on the definition of the motion picture image is well known. In microscopic work it is very objectionable. In the non-intermittent projector this objectionable feature is avoided because the mechanism has the advantage of operating without vibration.

Graininess, vibration, and errors in registration, are all sources of impairment of the motion picture image, and must be counted among the aberrations. These aberrations still exist and remain a problem in the motion picture image; and optical elements appear to be the only possible means that can be employed to reach and correct them.

Finally, it can be said, the most significant advantage is that the non-intermittent projector lets down the bars and opens the way to new and greener pastures. For almost within reach, are electrical and photo-electrical means of synchronization, a greatly extended range in rates of projection, higher resolving powers, and larger picture areas.

On the Utilization of Complex Optical Constants

FRANK MATOSSI
Naval Ordnance Laboratory, White Oak, Silver Spring, Maryland
April 6, 1949

RECENTLY Arzeliès[1] has warned against the indiscriminate utilization of complex optical constants in computing the properties of absorbing optical systems from formulas which do not include absorption. He expressly says that nothing justifies *a priori* the usual method of substituting complex constants into the formulas; only *a posteriori* can be determined whether the procedure is permitted or not. He does not give a mathematical reason for his objection but shows that in some instances there is indeed a difference between the results of his "direct method" and the classical one. If his statement is true, it would invalidate much of the work done on the optical properties of systems of absorbing layers. But, fortunately, it does not seem to be true.

The difference of the two procedures can be recognized not only in some particular cases but already in the statement of the refraction law. If i and r are the angles of incidence and refraction, respectively, the refraction law, according to the classical procedure[2] (subscript c) is given by

$$(\sin i/\sin r)_c = n_i,$$

where

$$n_i^2 = (\tfrac{1}{2})\{n^2-k^2+[(n^2+k^2)^2+4n^2k^2(gr]^{\frac{1}{2}}\},$$

n and $-k$ being the well-known real and imaginary parts of the complex refractive index $\eta = n - ik$. But according to Arzeliès, the refraction law is (subscript a) (cf. §11 of his second paper), in usual notation, $(\sin i/\sin r)_a = n n_i/(n_i^2 - k_i^2/\cos^2 r)^{\frac{1}{2}}$ with $k_i = nk/n_i$ or $n_i^2 - k_i^2/\cos^2 r = n^2 - k^2$. It can easily be seen that the two refraction laws are different.

The reason for this is that in Arzeliès' fundamental equations (§7), the velocity of propagation V is conceived as a real magnitude while a consistent treatment should require a complex velocity $V+iW$, the real part of which is the phase velocity. If this is introduced, the following expression for V is obtained from his equations:

$$V_a^2 = \{(1-k_i^2/n_i^2\cos^2 r)(n^2-k^2)+4nkk_i/n_i+(n^2+k^2) \times[(1-k_i^2/n_i^2\cos^2 r)^2+4k_i^2/n_i^2]^{\frac{1}{2}}\}/2(n^2+k^2),$$

while Arzeliès has

$$V_a^2 = (1-k_i^2/n_i^2\cos^2 r)/n^2.$$

With V_a, the refraction law indeed is again $\sin i/\sin r = n_i$ as it ought to be.

Therefore, practically all results of Arzeliès' direct method, I regret to say, must be deemed erroneous and his criticism, inconclusive (as far as absorbing media are concerned). We need not abandon the utilization of complex constants or rely on divination for the correct use of them, if there are no specific mathematical reasons prohibiting their use or expressly requiring a real velocity in spite of a complex index. But since the optical properties are governed by Maxwell's equations (with boundary conditions) which admit solutions with a complex refractive index and the usual meaning of its real and imaginary parts, there is a strong reason for, but none against, the classical method.

[1] H. Arzeliès, Ann. de physique 1, 5 (1946); 2, 133 (1947).
[2] This is presented most clearly and straightforwardly in W. Koenig's article in *Handbuch der Physik*, Vol. XX, Chap. 6, Sections 40, 47, and 57.